SPRINGER NATURE
Reference

Xizheng Ke

Handbook of Optical Wireless Communication

Springer

Handbook of Optical Wireless Communication

Xizheng Ke

Handbook of Optical Wireless Communication

With 1048 Figures and 142 Tables

Springer

Xizheng Ke
School of Automation and Information Engineering
Xi'an University of Technology
Xi'an, Shaanxi, China

ISBN 978-981-97-1521-3 ISBN 978-981-97-1522-0 (eBook)
https://doi.org/10.1007/978-981-97-1522-0

© Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Preface

Optical Communication: From Wired to Wireless

Transportation, energy, and communication are the three basic structures of human society. Transportation is the transfer of people and materials, energy is the transmission of energy, and communication is the transmission of information. All three greatly affect China's national economy.

The relationship between China's transportation, economic, and social development is undergoing tremendous changes, promoted by the application of modern information technology and innovative business models. The trend of integration, interaction, and mutual promotion is obvious. It also profoundly changes the traditional production, life, and organization of transactions, and has become an important force to drive the economic growth and social development of old and new systems.

Modern energy systems are based on a multi-energy structure that combines renewable energy with gaseous energy. Relying on smart energy technology coupled with clean energy and the Internet, the traditional energy system has gradually evolved to a new energy system, which is a necessary stage for future energy systems.

The deep integration and interdependence of transportation, energy, and information technology is a historical necessity. The tide of informationization, featuring digitization, networking, and intellectualization, is in the ascendant. Information technology is thriving daily and is fully integrated into the production and lives of human society. It is constantly interfusing with various industries, and is affecting the world's economic, cultural, security, and competition patterns.

Optical communication includes both wired and wireless methods. Cable optical communication requires a fixed medium. Optical wireless communication does not require a fixed medium and includes space laser, ultraviolet (UV), visible, and underwater blue-green communication, among others.

Fiber Communication

Optical fiber is the cornerstone of our digital society. Optical communication is a huge system, and optical transmission networks are the basis of the entire digital society. Following are some future development trends of optical communication technology.

1. The main task of optical communication is to transmit data and continuously expand bandwidth and capacity. In addition to improving its single-wave capacity, it should further expand its spectrum bandwidth.
2. Light is an energy-saving technology. The higher the optical-domain proportion in a transmission network, the lower the overall energy consumption. Operators should consider the top energy issues, such as reducing the overall network energy consumption, reducing the complexity of the network operation and maintenance, and reducing the amount of human intervention. From the optical-communication perspective, the transmission-energy consumption per unit bit kilometer and unit-bit switching-energy consumption should be further improved.
3. The changing role of metropolitan area networks should also be considered. A metropolitan all-optical network contains three layers: the metropolitan core, convergence, and access. A metropolitan all-optical network and the cloud network are closely related to each other. The metropolitan network not only intersects with data centers, but also enables operators with handheld devices to enter cloud business relations with government and enterprise customers. For example, operators can match private-line cloud services by providing optical broadband access.
4. Intelligent operation and maintenance are also necessary. After management is centralized, operators can intelligently operate the entire transmission network by introducing artificial intelligence (AI) and big-data technologies.

Optical Wireless Communication

Optical wireless communication does not require a fixed medium. Moreover, it is expected to become the main communication mode of satellite Internet, because of its high bandwidth, high security, and small device size. The main technical bottlenecks of optical wireless communication are narrow-beam long-distance divergence and atmospheric attenuation. The main technologies are automatic target acquisition and tracking, communication transceivers, atmospheric-turbulence compensation, integrated optical-electromechanical design, tethered-balloon relays, and stratospheric-airship relays.

Optical wireless communication also belongs to the field of applied optics, which combines optical, precision mechanical, electronic, material, computer, and other technologies to solve a variety of engineering application problems. It focuses on the acquisition, transmission, processing, storage, and display of light information and

the application of laser technology. It is directed toward multi-frequency bands, such as ultraviolet, visible, infrared, intensity, spectrum, polarization, and phase, as well as other dimensions, such as detection, imaging, and ranging. New communication and other multi-functional technologies are being developed.

Light has a much higher frequency than microwaves do, and as a communication carrier, it has a wider utilization band. At present, the working-frequency band of optical wireless communication is mainly 365 THz–326 THz (the corresponding wavelength range is more than 820 nm–920 nm), and there is no radio-frequency interference between devices. Waves, as information carriers, can transmit data rates up to 10 Gbit/s, and methods using wavelength-division multiplexing (WDM) technology can further improve this.

Optical wireless communication technology combines the advantages of radio and optical communication, and uses light as the communication carrier. Optical wireless communication technology has the advantages of a strong anti-jamming ability, high security, high communication rate, fast transmission speed, convenient band selection, and large information capacity. It has a small size, light weight, low power consumption, simple construction, and good flexibility and mobility. It has great strategic and application value in military and civil fields. In addition, the laser's advantages of a short wavelength, strong penetration, good directionality, and energy concentration also reduce the aperture of the receiving telescope, and eliminate the huge dish antenna of the radio communication system. The receiving system can also be made smaller and lighter. Thus, the construction and maintenance of laser communication-related systems and equipment are relatively inexpensive.

Optical wireless communication faces a dilemma. The reliability of light transmissions is challenged by such factors as atmospheric absorption and scattering, atmospheric turbulence, and background light. The main factors affecting satellite–ground light communication are atmospheric absorption and scattering, atmospheric turbulence, background light, clouds, rain, etc.

Atmospheric laser absorption is mainly due to the interaction between atmospheric molecules and a laser of a specific wavelength. Atmospheric scattering is also related to the wavelength, and can be described by the Beer–Lambert law. Atmospheric turbulence is caused by changes in atmospheric temperature and pressure, which redistribute the signal strength and phase, i.e., wavefront distortion. It is usually described using a gamma–gamma distribution or a more accurate two-parameter Weibull distribution. The influence of background light comes from the sun, other stars, and scattered light. Clouds and rain affect laser communication mainly from the angles of absorption and scattering.

Trends in Optical Wireless Communication

The technological trends of optical wireless communication are developing toward high speeds, networks, multipurpose applications, and integration. All-in-one systems, full photoluminescence, and multispectral communication are also discussed.

High Data Rates

With the rapid development of high-speed modem and transmission technology for space laser communications, future satellite-to-Earth optical wireless communication link rates are expected to reach 100 Gbit/s. High-speed laser communication uses high-order modulation methods, such as orthogonal phase-shift keying, orthogonal amplitude modulation, and multiplexing, such as wavelength-division, time-division, and orbital-angular-momentum multiplexing. Short-distance rates can reach Tbit/s magnitudes.

Networks

With the development of globalization and information technology, there is an urgent need to build spatial networks that do not depend on ground networks, with seamless global coverage, high bandwidth, and high performance. Therefore, space-based broadband-transmission networks are an important trend in the future.

Optical wireless communication technology is gradually moving from point-to-point modes to relay-forwarding and building optical wireless networks. The main difficulties when constructing laser networks are the small laser divergence angle, dynamic access of optical signals, and large impact of a space environment. Consequently, it is necessary to break through the one-to-many problem of laser-communication technology, study dynamic routing to solve the access scheme, and search for a joint laser–microwave communication system. Because of its simple optical principle, using a rotating parabolic structure to design one-to-many point optical antennas for laser-communication networking between multiple satellites is a major breakthrough in exploring and solving this difficult problem.

Multipurpose Applications

With the maturing of optical wireless communication technology, its advantages of a high modulation rate, long transmission distance, and low energy consumption have become increasingly prominent. At present, optical wireless communication technology has been widely used in the broadband data transmissions of inter-satellite, sky, space, and ground links, and has gradually expanded to deep-space detection, underwater and ground-access communications, etc.

Deep-space exploration involves the exploration of the moon, distant objects, or space. It is an inevitable step in understanding the solar system and the universe, discovering the origin and evolution of the universe, and expanding the living space of human beings. The implementation of a lunar-exploration project opened the curtain of deep-space exploration in China, and was followed by a Mars-exploration project.

Underwater optical wireless communication has the advantages of a large capacity, high bandwidth, good confidentiality, and strong anti-jamming ability. It has

become an important communication technology developed by the world's major powers. In addition, space laser communication technology can play an irreplaceable role in applications where optical cables cannot be laid, such as between islands, urban buildings, complex field environments, etc.

Optical wireless transmission technology using visible light for data communication combines the functions of lighting, communication, and control positioning. It is easy to integrate with existing basic lighting facilities, and conforms to national energy-saving and emission-reduction strategies. It is expected to gradually become the main broadband transmission method for ultra–high speed universal optical networking in the future intelligent era.

Integration

Because of the advantages of lasers in high-speed communication and accurate ranging, the integration of laser ranging and communication technology has received increasing attention in recent years. Integrated designs of laser ranging and communication are based on high-speed communication and enables precision ranging. The same laser beam and hardware platform are used for ranging and information transmissions, thus achieving the dual functions of ranging and communication with the same set of devices.

In 2013, the United States National Aeronautics and Space Administration's (NASA's) liquid-crystal display (LCD) system successfully implemented on-track demonstration verification of monthly high-speed laser communications and high-precision ranging, with the ranging accuracy reaching 3 cm. In 2006, scientists at Pennsylvania State University conducted an evaluation study of a laser/microwave hybrid transmission. It was found that the laser link was strongly affected by clouds, mainly due to the attenuation and scattering of cloud particles; however, the availability of the overall link can be greatly improved by introducing microwave links.

Usually, in a laser/microwave hybrid transmission mode, the laser link can pass through the laser transmission; however, the microwave link cannot be used to transfer the data. This method not only inefficiently uses the overall channel bandwidth, but also causes unnecessary transmission interruptions when switching between laser and microwave links.

All-in-One Systems

Space laser communication and ranging have many similarities; for example, they both require pointing, acquisition, and tracking (PAT) units, pulse-timing detection, and have advantages of high-speed communication and precision ranging. Therefore, they can be combined to form a composite communication and ranging system by sending and receiving a single laser beam, i.e., integrated laser communication and ranging.

For example, NASA's lunar laser-communication demonstration system has been successfully implemented to verify lunar–terrestrial high-speed laser communication and high-precision ranging in orbit. The Beijing Institute of Telemetry Technology has also completed an integrated high-speed communication and ranging design, based on coherent communication.

Integrating laser communication and ranging produces a single device with multiple functions, thus reducing the size and power-consumption requirements. However, it has disadvantages, such as poor anti-jamming capability and weak ranging optical energy. Fortunately, these can be solved by adopting technologies, such as common ranging and communication wavelengths, dual modulation systems, and pseudocode and communication-signal conversion.

Full Photoluminescence

At present, a satellite laser-communication relay node adopts an "electric relay" mode, which converts the weak received optical signal into an electric signal. It then amplifies the signal, recovers the clock, and filters and denoises the signal in the electric domain before modulating it to the laser carrier and transmitting it out. Thus, there is a process of light to electricity/electronics to light for this signal. The "electric relay" mode mainly suffers from a complex system structure (back-to-back transmitters and receivers are required), electronic bottlenecks (the electronic chip rate is limited to Gbit/s), etc.

The European Space Agency's (ESA's) high-throughput optical network plan proposes to build a high-throughput all-optical network by planning two satellites, GEO and LEO, to be used as all-optical link nodes for in-orbit demonstrations and user data-collection validation and distribution. The all-optical network is a broadband network that replaces traditional electrical nodes with optical nodes and connects them into a network with laser links; this enables information to be transmitted and processed entirely in the optical domain. The all-optical network uses advanced all-optical technologies, such as transmission, amplification, regeneration, and switching. To realize an all-optical laser-communication network, it is necessary to break through in key technologies, such as on-star all-optical relay-regeneration and switching technologies.

Multispectral Communication

Laser technology can be applied to a variety of spectral bands, from ultraviolet to infrared, and even terahertz. For example, NASA has begun to apply a terabit-class infrared-transmission payload to perform SpaceX's Transporter-5 carpool mission.

Each spectrum has different cloud-penetration abilities and suffers different levels of anti-electromagnetic interference. For example, the ultraviolet and visible-light bands have poor cloud-penetration and high anti-electromagnetic interference abilities. The terahertz spectrum has a strong cloud-penetration ability, second only to

radio, but its anti-interference ability is better than radio. The infrared light band's cloud-penetration ability is worse than terahertz, but its anti-electromagnetic interference ability is stronger.

Therefore, future space–laser communication will vigorously develop the UV, visible light, mid-infrared, terahertz, and other multispectral communication-mode combinations to take advantage of the different strengths of the spectral communication systems. Moreover, in-depth research on multispectral communication will be emphasized, with a view to achieving uninterrupted and barrier-free communication under different environmental conditions, through the joint application of multiple spectra.

Beaconless Light

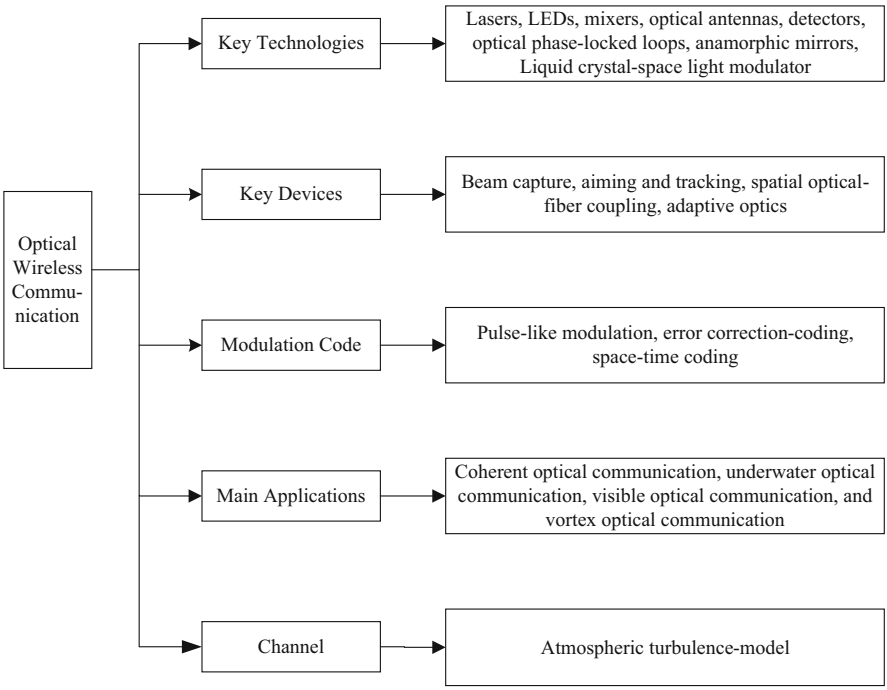
An optical wireless communication system typically uses two wavelengths of light: one wavelength to maintain alignment between the two terminals, and another wavelength for the transmission of information. Two lasers have different transmit power, modulation rate, and light dispersion angle, increasing the complexity of the system design. The direction of future development is not to beacon light, directly using the signal light for alignment.

Non-Co-Optical Axis

The transmitter and receiver of Optical wireless communication require a common visual axis. However, when the communication distance is far away, due to the disturbance of the wireless channel, the light beam will produce off-target, resulting in communication interruption. By adopting the non-co-axial technique, the transmitter and receiver are adjusted individually, which suppresses the effect of the disturbance of the wireless channel on the communication system and also reduces the requirement for ATP technology.

Structure of This Book

The figure in the following page shows a structural block diagram of the main contents of this book. This book analyzes and discusses the key technologies of optical wireless communication, including beam capture, aiming and tracking, spatial optical-fiber coupling, and adaptive optics. It also discusses the research progress and control algorithms of key devices used in optical communication systems, such as lasers, LEDs, mixers, optical antennas, detectors, optical phase-locked loops, deformation mirrors, and liquid-crystal spatial light modulators. The research progress of modulation and error coding in optical wireless communication



Block diagram of the main contents of the book

is summarized, and several typical optical wireless communication methods are introduced, such as coherent optical communication, underwater optical communication, visible optical communication, and vortex optical communication. The atmospheric-turbulence model of an optical wireless communication channel is introduced. Research progress and development trends are also discussed.

Contents

Volume 1

1 Optical Wireless Communication 1

2 Research Progress on Satellite Laser Communication Networks 21

3 Pulse-Like Position Modulation Technology 45

4 Communication Lasers and Their Modulation Technology 85

5 Research Progress on Passive Modulation in Free-Space Optical Communication 129

6 Detectors and Their Noise Models 157

7 Noise Models in Optical Wireless Communication Systems 205

8 Adaptive-Threshold Detection Technology 237

9 Four-Quadrant Detector Light-Spot Detection Principle 269

10 Optical-Antenna Technologies 297

11 Research Progress on One-to-Many Transmitting Antennas for Optical Wireless Communication 321

12 Acquisition, Pointing, and Tracking 357

13 Spatial Optical-Fiber Coupling Technology 393

14 Atmospheric-Turbulence Models 441

15 Atmospheric-Turbulence Suppression Methods in Optical Wireless Communication 487

16 Propagation of Partially Coherent Beams in Atmospheric Turbulence 517

17 Progress in Research on Channel Equalization in Optical Wireless Communication 559

18 Error-Correction Coding 581
19 Optical Wireless MIMO Technology and Space–Time Coding . . . 615
20 Space–Time Coding 631
21 Experimental Study on Optical Wireless Coherent Communication 693
22 Factors Affecting Coherent-Detection Sensitivity 799
23 Adaptive Optics Technology 829
24 Mode Methods in Adaptive Optics 875
25 Research on a Decoupling Algorithm for a Dual-Deformable-Mirror Wavefront-Correction System 917

Volume 2

26 Optical Phase-Locked Loops 941
27 Deformable Mirrors and Their Control Algorithms 987
28 Liquid-Crystal Spatial Light Modulators and Their Applications 1015
29 Mixers 1065
30 Principles and Development of Optical Amplifiers 1129
31 Key Technologies in Underwater Optical Wireless Communication 1153
32 Experimental Study on Underwater Optical Wireless Channel Models 1213
33 Principles and Research Progress on LEDs 1257
34 Principle and Research Progress of Visible Light Communication 1307
35 Indoor Visible-Light Communication and Its Heterogeneous Fusion Networks 1347
36 Optical-Camera Communication 1383
37 Indoor Visible Light Positioning Technology 1413
38 Research Progress on Visible-Light Communication Uplinks 1435
39 Research Progress on Indoor Visible Light-Source Layouts 1463
40 Resource-Allocation Strategy for VLC/RF Systems 1495

41 Indoor Visible-Light Location Based on a Fusion Clustering Algorithm 1527

42 Experimental Study on Chromaticity Control in Visible-Light Communication Systems 1549

43 Visible-Light Channel-Equalization Algorithm Based on the Fusion of a Neural Network and a Temporal-Feature Memory Structure 1583

44 Ultraviolet Non-Line-of-Sight Communication 1619

45 Experimental Study on Ultraviolet Optical Communication Noise 1647

46 Research Progress on OWC/RF Hybrid Communication Systems 1703

47 Performance Analysis of a Hybrid FSO/RF/THz Relay Communication System 1729

48 Orbital Angular Momentum Beam Techniques 1755

49 Research Progress on Aircraft Relay Optical Wireless Communication 1829

50 Research Progress on Optical Wireless Communication in Industrial Internets 1865

About the Author



Xizheng Ke is Doctor of Science, Second-Level Professor, Foreign academician of the Russian Academy of Natural Sciences, Head of the Department of Information and Communication Systems at Xi'an University of Technology, and Deputy Director of the Key Laboratory of Military-Civilian Joint Construction of Intelligent Cooperative Networks in Shaanxi Province. He is a famous teacher in Shaanxi Province, Fellow of Chinese Institute of Electronics, Director of Chinese Optical Engineering Society, Executive Director of Shaanxi Optical Society. He received a Bachelor's degree from Shaanxi Institute of Technology in 1983 and a Doctor of Science degree from the University of Chinese Academy of Sciences in 1996. From 1997 to 2002, he conducted post-doctoral research at Xidian University and the Second Artillery Engineering College. Dr Xizheng Ke is a member of the editorial board of *Journal of Electronics*, *Infrared and Laser Engineering*, *Journal of Electronic Measurement and Instrument*, *Laser Technology*, *Applied Optics*, *Journal of Xi'an University of Technology*, *Journal of Time and Frequency*, and *Journal of Atmospheric Science Research*. He is the national expert for reviewing science and technology awards and a member of the disciplinary review group of the Shaanxi Academic Degrees Committee. In 2000, Dr Xizheng Ke won the Outstanding Young Scholar Award of the Chinese Academy of Sciences. In 2009, he was awarded the title of "Outstanding Science and Technology Commissioner of Guangdong Province" by the Ministry of Science and Technology. In 2015, he was awarded the title of "A Golden Phoenix on Green Poplar" in Yangzhou City. In 2018, he won the China Industry-University-Research Innovation Award,

and in 2019, he won the China Industry-University-Research Innovation Achievement Award. Since 2001, he has won 16 provincial and ministerial science and technology awards, including 1 first prize and 5 second prizes. He has obtained more than 20 nationally approved invention patents, published 9 monographs in Science Press, and published more than 400 academic papers in domestic and foreign journals. More than 30 PhD students have been trained under his direction.

Optical Wireless Communication

1

Contents

Introduction 2
Optical Wireless Communication Development Status 3
Foreign Development Status 3
Domestic Development Status 9
Optical Wireless Communication Systems 11
Components of Optical Wireless Communication Systems 11
Features of Optical Wireless Communication Systems 13
Problems of Optical Wireless Communication Systems 14
Key Technologies 15
High-Quality Optical System Design 15
Fast, Accurate, and Stable Capture Alignment Tracking Technology 16
High-Rate, High Power-Emitting Lasers 16
Atmospheric Channel Impact Compensation Technology 17
Summary 17
References 18

Abstract

Optical wireless communication is an emerging wireless communication technology, which has faster transmission speed, higher bandwidth, and lower power consumption compared with traditional communication technology. This chapter systematically summarizes the development status, system composition, and key technologies of optical wireless communication at home and abroad. The key technologies and technical difficulties are analyzed, and the future development trend of space laser communication is summarized, which points out the direction for the future research and development of space laser communication.

Keywords

Optical wireless communication · Communication system · Key technology

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_1

Introduction

Compared with traditional radio-frequency (RF) communication, optical wireless communication has the advantages of high capacity, good confidentiality, light weight, low power consumption, strong anti-interference ability, and abundant frequency resources. Compared with fiber-optic communication, optical wireless communication can transmit under special link conditions, such as interstellar, star–ground, and air–ground, and can save expensive fiber-optic laying and maintenance costs.

These advantages make space optical communication technology an important research direction for a new generation of information and communication technology. Space laser communication, ultraviolet (UV) nondirect vision communication, visible-light communication, and blue–green light underwater communication are all important application scenarios for optical wireless communication.

Optical wireless communication includes two systems: intensity modulation/direct detection, which can also be called direct detection, and higher-order modulation/coherent detection, which can also be called coherent optical communication. The higher-order detection uses a local oscillating light source at the receiver side to coherently demodulate the received signal.

In terms of modulation and demodulation, coherent optical communication systems have many advantages over direct-detection systems, such as high sensitivity, a longer communication distance, many different modulation/demodulation methods, and the abilities to filter out very strong background light, select different light wavelengths for communication without significant wavelength restrictions, and recover damaged signals through digital signal processing algorithms.

Although optical wireless communication technology has many advantages, its disadvantages somewhat limit its application prospects. In an optical wireless communication system, the atmospheric channel characteristics can seriously interfere with the propagation of a signal beam containing information. This can cause light-intensity flicker, spot drift, beam expansion, phase undulation, pulse spreading, beam bending, and drift. These seriously reduce the uniformity and stability of the received optical signal. In addition, scattering and absorption will weaken the beam energy.

With increased communication distance, a strong turbulence environment, large optical system aperture reception, and other conditions, the laser output power gradually increases, the laser-cavity mirror suffers from thermal deformation, and the beam wavefront undergoes phase distortion, which reduces the beam quality. This directly leads to an increased communication error rate, is not conducive to establishing a stable communication link, and in serious cases, will interrupt communications.

To reduce the abovementioned atmospheric effects on wireless coherent optical communication and improve the system performance, corresponding measures must be taken to suppress the atmospheric effects when designing the system, so that the optical wireless communication can adapt to the long-distance, high-speed, and low transmission bit-error rate (BER) requirements.

Research on atmospheric turbulence suppression and performance compensation techniques has been conducted, and various methods to suppress atmospheric turbulence effects on wireless coherent optical communication systems have been proposed. These include large-aperture reception, diversity transceivers, partially coherent optical transmissions, coding and modulation, time-domain and threshold optimization, and adaptive-optics techniques.

Optical Wireless Communication Development Status

Foreign Development Status

In 1880, Alexander Graham Bell, a famous American inventor, developed the "optical telephone," which was the first time light was used as a carrier to send information, using the atmosphere as a transmission medium. This is regarded as the beginning of modern optical wireless communication [1]. After the development of the first solid-state ruby laser, the spectrum, coherence, directionality, and other light-source characteristics were improved, and research on optical wireless communication gradually entered the experimental stage [2]. However, owing to the limitations of early device performance and external factors, such as atmospheric channels, research on optical wireless communication was stagnant for a long time, while fiber-optic communication technology flourished and was widely used.

In the 1970s, as technologies were updated, optical wireless communication again became the focus of worldwide attention, and various military organizations tried to apply it to their equipment. The military and aerospace institutions of Europe, the United States, and other advanced countries successfully conducted numerous experiments on optical wireless communication transmissions.

In the 1980s, optical wireless communication technology was combined with wavelength-division multiplexing (WDM) access, optical amplification, and adaptive optics, which greatly improved optical wireless communication in terms of the transmission distance, transmission capacity, reliability, etc. In the late 1980s, some manufacturers began to develop commercial optical wireless communication technology, solving a series of key problems, and further reducing the cost. By the 1990s, optical wireless communication technology was widely used in civilian products.

In the late twentieth century, early major twenty-first-century communications companies began to research and apply optical wireless communication technology on a large scale. In recent years, with the continuous development of erbium-doped fiber amplifiers, adaptive optics, coherent detection, and other technologies, the transmission distance and reliability of wireless laser communication have greatly improved and it has re-emerged as a hot research topic.

United States

The United States was the first country to study free-space optical communication technology. Since the 1960s, it has conducted numerous experiments on optical communication under various scenarios. As early as the mid-1980s, the

Massachusetts Institute of Technology's (MIT's) Lincoln Laboratory built an experimental interplanetary laser-outlier coherent detection communication system using frequency-shift keying modulation, with modulation rates up to 220 Mbit/s. The performance of each module of the system was tested in a space environment. However, the project failed to conduct actual flight tests and only completed demonstration experiments on the ground [3].

In 1981, the US Air Force conducted a demonstration experiment between ground stations and aircraft, using a frequency-doubled yttrium–aluminum–garnet (YAG) laser with a 532-nm wavelength. The communication distance and rate reached 100 km and 1 Gbit/s, respectively. They also successfully demonstrated wireless light detection and ranging (LIDAR) [4].

In 1998, NASA's Jet Propulsion Laboratory (JPL) conducted a terminal system-level assembly and performance test of an optical communication demonstrator using a laser communication terminal measurement station [5]. Ground-based optical-link experiments were conducted at a distance of 4 km. The ground terminal used for the demonstration experiments was the Optical Communications Telescope Laboratory (OCTL); its construction was completed by the JPL in 2000. This laboratory was used to study the performance of multiple laser transmissions in space, allowing optical communication with both low and synchronous orbit terminals [6].

In February 1998, Lucent Technologies, Inc. built an experimental system for optical wireless communication with a rate of 10 Gbit/s for inter-near-earth links. Because the signal-light transmission was in the atmosphere, the communication distance was limited by the atmosphere's complex environment; the thermal expansion and contraction of buildings also affected the collimation of the beam. The Astro Terra company added an automatic tracking system to correct the influence of buildings, thus maintaining the beam collimation. In August of the same year, Lucent and Astro Terra successfully tested a prototype of a optical wireless communication system with a 1.5-km link distance and 1.5-Gbit/s data rate [7].

In 2000, the Tri-Service Experiments 5 (TSX-5) satellite was successfully launched. It carried an on-board optical communication terminal; however, the position and attitude-control accuracy of the satellite was not considered in the design. Consequently, the on-board optical communication terminal was never able to capture and track the signal light emitted from the optical ground station, despite a dozen link attempts [8].

In 2008, the United States and the German Space Agency (DLR) initiated an international collaboration to explore the capability of high-rate coherent laser communication through the atmosphere, establish a communication link, and test the performance of terrestrial laser communication terminals. In the same year, NASA launched the Deep Space Optical Communications Program, which included the Lunar Laser Communications Demonstration. Its aim was to verify the feasibility of laser communications between Earth and the moon, using small, lightweight communication terminals [9].

By 2013, the MIT Lincoln Laboratory had developed an onboard optical communication terminal. On September 6, 2013, the successful liftoff of the terminal

aboard the Lunar Atmosphere and Dust Environment Explorer spacecraft marked the beginning of NASA's Lunar Laser-Communications Demonstration Validation Program in-orbit experiment. The program was designed to validate two-way laser communications between the on-board optical communication terminal and an optical ground station. This was the world's first laser communication program at a lunar–terrestrial distance.

Considering the weather, three optical ground stations communicated with the onboard optical communication terminal: one primary and two backups. These included the terrestrial optical communication terminal at White Sands, New Mexico, the Lunar Laser Communication Telescope Laboratory terminal at the Table Mountain Facility in California, and the Lunar Optical Ground System, the European optical ground system for laser communication at Tenerife, Canary Islands [10]. In October 2013, the first bidirectional lunar–terrestrial laser communication was successfully carried out between an onboard optical communication terminal and a terrestrial optical communication terminal. It had an uplink data transmission rate of up to 20 Mbit/s and a downlink data transmission rate of up to 622 Mbit/s.

In 2017, the Deep-Space Optical Communications program returned data from Mars at a rate higher than 250 Mbit/s over a communication distance of approximately 630-million km, with a 28-kg terminal mass and 76 W of power [11].

The primary optical ground station for the Lunar Laser Communication Demonstration program was designed and built by Lincoln Laboratory [12]. The optical antenna of this ground station had a space diversity design that suppressed the effects of atmospheric turbulence on the uplink and downlink transmissions; it was scalable, cost-saving, and easily movable. The optical antenna, which included a glass-fiber temperature-controlled housing, was approximately 4.5 m high and weighed approximately 7 t. This is a reduction of approximately 75% in size and weight, compared to conventional radio ground antennas.

The optical antenna contained four homemade refractive-transmitting telescopes with 15-cm apertures, each connected to a transmitter in the control room via a single-mode fiber, which could send an uplink signal light with 10 W of power. In addition, it included four commercial Cassegrain reflecting receiving telescopes with 40-cm apertures. Each telescope coupled the received downlink signal light into a multimode fiber and detected it through an array of photon-counting superconducting nanowires in a control room [13].

On June 5, 2014, the NASA-JPL Optical Communications Telescope Laboratory (OCTL) optical ground station successfully received a 175-megabyte high-definition video called "Hello, World" from an optical communications terminal installed on the International Space Station (ISS) in 3.5 s. This was a data rate equivalent to 50 Mbit/s and marked the success of the Optical Communications Telescope Ground Station demonstration experiment. The Optical Communications Telescope Ground Observatory is a part of the Jet Propulsion Laboratory's Phaeton program. Owing to cost constraints, the Optical Communications Telescope Ground Observatory was designed and built by the JPL using commercial off-the-shelf components, with a design mission run time of 90 days. The project was designed to prepare the ground station for optical communications and low-orbit satellite experiments [14].

The successful liftoff of two 1.5-U CubeSats by NASA in 2017 marked the implementation of the Optical Communications and Sensors Demonstration Validation Program in-orbit experiment. It was designed to validate high-speed laser communications between CubeSats and with optical ground stations. The mission successfully conducted laser communications between CubeSats and optical ground stations in August 2018, with uplink data transfer rates of up to 10 kbit/s and downlink data transfer rates of up to 200 Mbit/s.

The US Marine Corps tested a tactical line-of-sight optical communications network at US Army Camp Hansen in Okinawa, Japan, on August 21, 2018, to observe its performance in a wet, variable weather environment. In 2013, the tactical line-of-sight optical communication network demonstrated a 100 Mbit/s high-resolution image- and video-transmission capability of over 50 km in evaluation tests [15].

In September 2021, Alphabet, Google's parent company, opened the Taara research project to develop an atmospheric laser communication system to establish a laser communication link with a 20-Gbit/s communication rate and 5-km communication distance between the cities of Brazzaville and Kinshasa in the Democratic Republic of Congo. It transmitted 700 TB of data over broadband Internet in 20 days [16].

In October 2023, NASA plans to conduct an exploration mission to Psyche, an asteroid located between Mars and Jupiter whose main components are iron, nickel, and rare metals (gold, platinum, copper, etc.). The mission will include a deep-space optical communication terminal and a technology demonstration of the Deep-Space Optical Communication project [17]. One of the two main objectives of the Deep-Space Optical Communications project is to validate the high photon-efficiency reception technology recommended by the Space Data Systems Advisory Committee [18]. This standard uses the technical solution of a photon-counting receiver and a high-peak-power transmitter that can deliver several bits of information per detected photon. The second objective is to demonstrate acquisition, tracking, and beam-pointing control techniques in deep-space links.

The Deep-Space Optical Communications project consists primarily of the following:

1. A vehicle laser transceiver capable of receiving a 1064-nm uplink/beacon laser and transmitting a 1550-nm downlink laser.
2. A ground-based laser transmitter (GLT) located at the Optical Communications Telescope Laboratory in California to transmit 1064-nm uplink/beacon lasers.
3. The Hale Telescope at Palomar Mountain, California, which was used as a ground-based laser receiver with a tungsten-silicide superconducting nanowire single-photon detector array with photon-counting reception capability.
4. The Mission Operations Center, which is responsible for overall coordination.

The downlink uses a serial cascaded pulse position modulation coding method with a 4-W transmit power and communication rates of approximately 132 Mbit/s and 2 Mbit/s at distances of 0.25 AU and 2 AU, respectively (one astronomical unit

(AU) is the average distance from the Earth to the Sun, which is approximately 150 million km.). The uplink uses low-density parity check coding and a communication rate of approximately 1.6 kbit/s at a 1-AU distance.

Europe

In the late 1970s, the European Space Agency researched and developed high-speed laser signal modulation technology and evaluated the corresponding modulators. This event was considered the formal entry of the European Space Agency into the field of free-space optical communications. Subsequently, the success of the European Space Agency's Semiconductor Laser Inter-satellite Link Experiment (SILEX) and the German laser communication terminal on the TerraSAR–Xsatellite (LCTSX) marked the beginning of Europe's development in the field of free-space optical communications. The success of the German LCTSX project marked a new era in the development of optical communications in Europe [19].

From 2003 to 2009, the European Space Agency conducted up to 393 two-way free-space optical communications between the Earth and stars, using the ARTEMIS satellite and optical ground stations. Experimental statistics show that its communication success rate reached 91%, and its total communication time reached 78 h [20].

In 2002, the DLR started the LCTSX program, which aimed to conduct satellite–ground coherent laser communication experiments, with two sets of laser communication terminals using binary phase-shift keying/coherent detection. The terminals were mainly developed by Tesat–Spacecom. The company used a neodymium-doped (Nd):YAG solid-state laser with a frequency stability of 50 MHz per day, linewidth of approximately 5 kHz, and 1064-nm operating wavelength, as the main light source in its development.

Subsequently, in 2005, the company conducted a spatial optical communication test at a rate of 5.625 Gbit/s and communication distance of 142 km on the islands of Tenerife and Palma, Spain. This experiment verified the performance of a laser communication system based on binary phase-shift keying and zero difference detection in a near-surface atmospheric environment and provided valuable experience for subsequent optical communication between satellites [21].

After completing these experiments, the researchers conducted a coherent optical communication experiment between the satellite and ground station with a communication distance of approximately 1000 km at Haleakala Volcano in 20010. The experiment was performed without the assistance of adaptive optics. Therefore, to reduce the impact of the wavefront distortion introduced by the atmosphere on the system, researchers minimized the receiving aperture of the ground station, while maintaining the receiving power.

During the course of the experiment, the researchers found that the BER of the satellite-to-ground downlink was relatively high, compared to the ground-to-satellite uplink. Finally, after analysis, it was unanimously concluded that this phenomenon was due to the limited transmit power of the transmitter carried by the satellite [22]. Thereafter, several improvements were made to the system in Germany, based on the experimental results, such as increasing the transmitting power and aperture. The improved system was named LCT-TDP1.

Owing to the success of the SILEX and LCTSX programs, at the end of 2008, the European Space Agency began to implement the European Data-Relay System (EDRS) program, which aimed to build a 1.8 Gbit/s space data highway between low-earth orbit (LEO) satellites and ground data centers through two synchronous orbit relay satellites (EDRS-A and EDRS-C). EDRS-A, an onboard optical communication terminal, was successfully lifted into its intended orbit on Eutelsat 9 B in January 2016, and EDRS-C was launched in August 20110. Prior to this, the synchronous orbiting satellite Alphasat and low-orbiting satellites Sentinel 1A and Sentinel 2A, the forerunners of the EDRS program, were launched between 2013 and 2015, respectively [23].

According to Airbus, as of May 18, 2018, the EDRS-A satellite had completed 10,000 laser connections with 910.8% reliability and had transmitted 500 TB of data. The European Space Agency planned to expand the EDRS program with the launch of another synchronous orbit relay satellite (EDRS-D) to become a data relay communication system with global coverage by 2020 [24].

In 2018, Germany used an airborne terminal, MLT-70-ATG, and an optical ground station, GS-200, in the FALCON project. Laser communication with a 10-Gbit/s communication rate and 10-km communication distance was achieved between the stratosphere and ground [25].

In 2019, Russia used the Artolink M1-GE-L model optical transceiver for building-to-building space laser communication experiments, with a communication wavelength in the 1550-nm band, link distance of 4.4 km, and rate of 1.25 Gbit/s. A design with a local static servo mechanism was selected. The weight of the optical transceiver was 9 kg, the system volume was 480 mm $\times$ 300 mm $\times$ 285 mm, and the power consumption was 46 W. In 2020, Russia used the improved Artolink M1-10GE model optical transceiver, and the communication rate increased to 10 Gbit/s with a 1.5-km communication link distance.

In 2021, Oxford University, UK, and Airbus Operations, UK, implemented reverse-modulated free-space laser communications between the ground and commercial unmanned aerial vehicles (UAVs). The system used an electrically absorbed modulator to modulate a backward reflector to establish a data link. The backward reflector had an effective aperture of 11 mm and a full field of view of 6.4°, achieving a 500-Mbit/s communication rate at a 560-m communication distance with a 8.6×10^{-4} BER [26].

Japan

In 1985, the National Space Development Agency (NASDA) of Japan began research on intersatellite laser communication from capture, tracking, and targeting technology, followed by system research and development. In 1986, Japan conducted a large number of research experiments with the United States [27].

Japan proposed the ETS-VI program in 1994 and ground stations successfully communicated with a satellite carrying a laser communication system developed under the program. Japan also proposed the Optical Inter-orbit Communication Engineering Test Satellite Program, which aimed to develop a miniaturized satellite test platform for optical communication in space [28]. In 2000, Japan developed a

two-way ultra-high-speed laser communication terminal to the Earth for use on the International Space Station. It was installed on an experimental platform called the "Kibo" loadable experiment module [29].

In 2006, the National Institute of Information and Communications Technology (NICT) of Japan conducted a two-way communication test with a ground station of the German Space Agency (DLR). The results showed that the BER of the data transmitted from the ground station to the Japanese satellite was only 10^{-7}, which once again verified the laser communication capability between the ground and LEO satellite.

Laser communication equipment also successfully achieved two-way laser communication with KODEN, an optical ground station located in Tokyo. This was the world's first laser communication experiment between a low-orbiting satellite and an optical ground station; an experimental study of uplink multibeam emission technology was also conducted [30].

A satellite-based coherent optical communication receiver, designed independently by the NICT of Japan, was successfully developed in 2009. The receiver adopted a zero difference coherent detection scheme and performed recovery operations on the carrier frequency and phase of the signal using a digital signal processing unit embedded in a field programmable gate array on board. The results showed that the receiver could demodulate binary phase-shift keying signals at a rate of 3 Gbit/s and orthogonal phase-shift keying signals at a rate of 6 Gbit/s [31].

Ten-Gbit/s bidirectional laser communication was achieved between ground stations and synchronous orbiting satellites in 2021. The related technology will be gradually applied to other aspects, such as space ranging and satellite networking. Japan will eventually achieve the goal of using its free-space optical communication technology to create a comprehensive information platform.

Domestic Development Status

Changchun University of Science and Technology has conducted laser communication experiments between the ground, ships, and aircraft. In 2007, a dynamic laser communication experiment with a horizontal channel between a ship and the ground was conducted. In 2011, two-way laser communication experiments were conducted, including a dynamic horizontal channel between two ships, a double dynamic oblique-range channel between an airship and a ship, and a double dynamic airborne horizontal channel between two helicopters.

In 2013, experiments were conducted on dual dynamic long-range and high-rate laser communication between two Y-12 aircraft. In 2015, a laser communication optical terminal between LEO-synchronous orbits was designed and developed, and in 2016, a laser communication optical terminal for synchronous orbit to ground communication was developed [32].

In 2011, a laser communication terminal was successfully launched onboard the OceanSat-2 satellite, which conducted a two-way laser communication test between

the satellite and ground base station. A laser link was established and stable communication was achieved [33].

The University of Electronic Science and Technology has researched intersatellite laser communication, satellite–ground laser communication, and other related fields. These include intersatellite optical communication capture and tracking technology, weak-beacon optical signal processing technology, satellite–ground coherent optical communication technology, and high-precision alignment information acquisition methods [34].

The Shanghai Institute of Optics and Precision Machinery, Chinese Academy of Sciences, has carried out a series of studies on the development of interplanetary and terrestrial laser communication coherent detection devices, optical phase-locking technology, and underwater optical communication [35].

Wuhan University realized laser communication between ships and shore buildings in 2013, and 1.5 Gbit/s high-speed laser communication between ships and shore buildings in 2018. The ship was able to maintain high-definition video communication while sailing [36].

The Air Force Engineering University has conducted research in the fields of space laser communication, star–ground laser communication, unmanned aerial laser communication, and visible-light communication [37].

From late 2016 to early 2017, the high-speed satellite–ground coherent optical communication terminal developed by the Shanghai Institute of Technology and Physics, Chinese Academy of Sciences, successfully carried the "Mozi" satellite to complete dual communication in orbit. The uplink used pulse position modulation with a 20-Mbit/s communication rate; the downlink used binary phase-shift keying modulation with a 5.12-Gbit/s communication rate [38].

Since 2014, the Institute of Optoelectronics Technology, Chinese Academy of Sciences, has developed a coherent optical communication system for long-range satellite–ground free-space communication and analyzed the improvement of its BER after adopting adaptive optical correction. In 2018, a ground station for satellite–ground laser communication was established. In addition, joint experiments were conducted on coherent laser communication between a synchronous orbiting satellite and the ground, by the Fifth Institute of Aerospace and the Institute of Optoelectronics. The laser communication terminal onboard the Practice 20 satellite platform was launched in the second half of 2019 [39].

The Anhui Institute of Optics and Precision Machinery, Chinese Academy of Sciences, has investigated light-wave transmissions in the atmosphere, including constructing observation fields for atmospheric-radiation parameters, and atmospheric-optical parameters. They also investigated the principles and methods of atmospheric-optical parameter measurement and developed key equipment. Long-term atmospheric radiation, optical parameter measurement, and database establishment were also studied [40].

In recent years, the Xi'an University of Technology has conducted in-depth theoretical exploration and experiments on the characteristics of laser transmissions in atmospheric turbulence, coding and decoding wireless laser communication systems, optical wireless subcarrier modulation, channel estimation and

equalization, and alignment tracking and capture techniques. It has also investigated coherent detection techniques, the design of optical transceiver systems, optical orthogonal frequency-division multiplexing modulation, optical multiple-input multiple-output systems, and optical wireless transmission theory and technology in complex environments.

In 2016, a space laser communication system terminal with an intensity modulation/direct detection method was successfully developed. It had standard Ethernet data access capability and could reliably transmit duplex data, voice, and images. Since 2017, the Xi'an University of Technology has successfully developed a differential binary phase-shift keying modulation communication terminal for an external differential coherent detection system, in cooperation with the China Academy of Space Research, with funding from the National Natural Science Foundation of China and major industrial chain projects in Shaanxi Province. It also began research on optical wireless communication network systems with ultra-long-range transmission and communication code rates (5–10 Gbit/s).

In 2018, the team developed a 10-km outfield experiment with an integrated transceiver and successfully transmitted video and voice signals. In 2021, the team initially developed a near-infrared-band space laser communication system with an intensity modulation/direct detection mode, with a maximum transmission rate of 10 Gbit/s. It used image tracking at the transmitter side and two-dimensional reflector control at the receiver side to suppress atmospheric turbulence. The system uses a non-co-axial operating system at the transmitter and receiver sides and achieves fast beam alignment with a 100-km link on the ground.

The development trends of space laser communication systems are in the direction of large bandwidth, high reliability, and beacon-free light and are moving from point-to-point transmission to point-to-multi-point transmission. Application scenarios are developing from initial space laser communications to terrestrial space laser communications. Moreover, 6G is also considered for application in space laser communication.

Space laser communication in atmospheric turbulence has been investigated, as well as traditional transmitters and receivers with centralized control of the common axis alignment tracking and capture technology, and transmitters and receivers with independent control of non-common axis technology development. These gradually reduce the dependence on a dynamic base on a stable platform.

Optical Wireless Communication Systems

Components of Optical Wireless Communication Systems

A complete wireless laser communication system is illustrated in Fig. 1.1. It includes an optical antenna (telescope), laser transceiver, signal processing unit, automatic follow-up system, and other components. The transmitter light source is a laser diode or light-emitting diode. After channel coding and modulation, the laser is used for

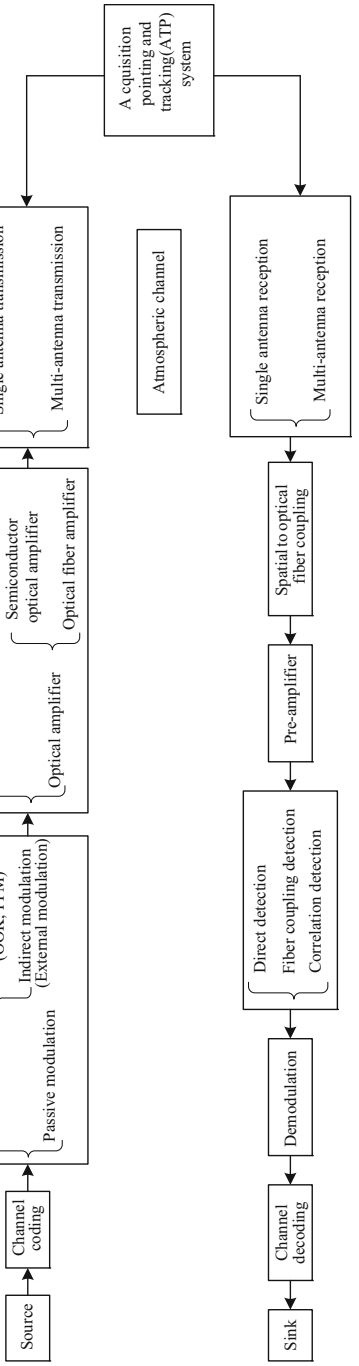


Fig. 1.1 Schematic diagram of a wireless laser communication system

direct output or amplified output by an optical amplifier. The beam is collimated by the transmitting antenna and transmitted through an atmospheric channel.

After being received by the optical antenna, the spatial light is coupled into a single-mode fiber by the optical-coupling module, amplified by the preamplifier, and then converted from an optical signal to an electrical signal by direct or coherent detection. It is then demodulated and decoded to finally complete the transmission from the source to the destination.

Features of Optical Wireless Communication Systems

Wide Communication Bandwidth and Large Information Capacity

Compared to microwaves, a laser carrier has a higher frequency and larger relative bandwidth, owing to its short wavelength, and its bandwidth is theoretically tens of thousands of times higher than microwave bandwidths. At present, transmission rates in fiber-optic communications have reached 1 Tbit/s; 100–622 Mbit/s terrestrial wireless laser communication systems have been practical. Demonstration experiments of a 5.65-Gbit/s wireless laser communication system have been conducted to verify the success.

Large Optical Gain and Small Laser Beam Dispersion Angle

When the antenna structure size is fixed, the shorter the wavelength, the greater the antenna gain, and the smaller the beam dispersion angle. Compared with microwaves, the wavelength of an optical wave is shorter; that is, the optical communication antenna has a higher gain. In addition, the laser beam dispersion angle is smaller than that of a microwave beam. Through the design of a moderate transmitting-antenna aperture, it is easy to achieve a laser beam dispersion angle on the order of microradians, which means better laser beam directionality and higher beam energy density. This feature makes the beam more conducive to long-distance transmissions in high-loss scenarios.

High Immunity to Electromagnetic Interference and High Confidentiality

Because the frequency band of a laser is far from the radio-frequency spectrum, laser communication is free from the interference of other electromagnetic waves in space. Its high confidentiality is reflected in the following:

1. The laser beam is extremely narrow and propagates in a straight direction; hence, it is not easily intercepted;
2. The laser communication beam is in the near-infrared band; this is non-visible light and also not easily detected;
3. Because the laser communication rate is higher, the transmission time is shorter under the same data-volume conditions; hence, it cannot easily be listened to.

Small Device Size, Light Weight, and Low Power Consumption

Because of the short wavelength of light waves (generally a few hundred nanometers to a few microns), according to the relationship between the antenna gain, beam angle, and antenna aperture, the wavelength year-on-year theorem states that the size and mass of an optical antenna in wireless laser communication are much smaller than those of an antenna in electromagnetic-wave communication. Meanwhile, the communication and beacon transmitter units generally use semiconductor lasers, which have low power consumption and high energy-conversion efficiency.

Problems of Optical Wireless Communication Systems

For the development of long-range, high-speed, and integrated practical wireless coherent optical communication technology, specific issues require further in-depth research.

Coherent Detection Must Satisfy More-Demanding Matching Conditions

Coherent optical communication requires a signal light and local oscillation light in the optical mixer to enable the mixing and balancing detector to convert the electrical signals. To meet the conditions of coherent detection, the signal-light and local oscillation-light linewidths must meet strict requirements. The signal light and local oscillation light must have the same wavefront phase, the polarization states must match, and the frequency difference must be within a certain range. The degree of matching between all factors directly affects the mixing efficiency of a wireless coherent optical communication system.

Matching the Complexity of the Adaptive-Optics Algorithm with the Atmospheric Environment

At present, most adaptive optics use the classical proportional–integral–differential control, and the time-varying characteristics of atmospheric turbulence limit traditional control algorithms. Intelligent control algorithms are used, typically based on the number of deformation-mirror drive units and wavefront sensor-array points, as well as the actual atmospheric environment, to make a comprehensive assessment and balance the practicality and control effect of the algorithm. The choice of the algorithm is based on the number of deformation-mirror drive units, wavefront sensor arrays, and the actual atmospheric environment.

Atmospheric Turbulence Reduces the Efficiency of Spatial Light-to-Fiber Coupling

A laser transmission through atmospheric turbulence reduces the spatial coherence of the beam, while the wavefront phase distortion caused by atmospheric turbulence significantly reduces the matching of the beam mode to the fiber mode for free-space transmissions. The coupling efficiency is reduced, which seriously affects the receiver sensitivity, BER, and other performance metrics.

Typically, single-mode fibers have a core diameter of less than 10 μm, which makes it difficult to couple spatial light into single-mode fibers in the presence of atmospheric turbulence. Multimode fibers can accommodate hundreds of single-mode fiber mode fields, because their maximum core diameter of over 61.5 μm allows for improved coupling efficiency with spatial light beams. However, multi-mode fibers are not compatible with existing single-mode fiber-based communication technologies, particularly in the field of coherent detection communications.

Atmospheric Turbulence Complicates Beam Alignment in Long-Range Optical Wireless Communication Systems

For point-to-point laser communication, a light beam with a divergence angle of approximately a micro-arc is used as the carrier for transmitting information. The prerequisite for establishing wireless laser communication is to maintain the precise alignment of the transceiver and transmitter view axes of both optical terminals. However, owing to the narrow beam and strong directional characteristics of a laser, the signal beam must be driven by the servo motor of a high-digit angle coder with extremely precise pointing; hence, it requires considerable time to complete the beam-capture process. This significantly reduces the practicality and efficiency of laser communication.

Key Technologies

High-Quality Optical System Design

Optical system design is one of the key technologies for space laser communication [41–44]. It is mainly reflected in the following three aspects.

Near Diffraction Limit Angle Optical System Design Technology

The beam generation quality can be mainly ensured by the integral beam expansion of the reflection light source and a two-stage telescopic system. This causes it to occur near the diffraction limit and is of great significance for enhancing the stability and safety of space laser communication.

Multi-optical Axis Consistency Mounting Technology

Laser communication systems generally contain five to seven optical axes. In the operation process, the multiple optical axes have high consistency requirements. This is mainly achieved through five-prism shift technology and assisted by charge-coupled device (CCD) subdivision detection technology to further ensure the efficiency of the application of this technology.

Optical Abutment Light and Small Design Technology

The main difficulty of this technology is that the system components are numerous, including dozens of optical components, lasers, acquisition, pointing, and tracking (APT) actuators, and transceiver components. Measures, such as laminated structure

design, modal analysis methods, and selecting lightweight, radiation-resistant materials, can be used to substantially improve the adaptability of space laser communications.

Fast, Accurate, and Stable Capture Alignment Tracking Technology

Highly accurate, precise, and stable acquisition, pointing, and tracking (APT) technology, for capturing alignment tracking under long-range conditions, is one of the most important technologies for space laser communication [45]. The specific implementation steps are as follows:

1. Determine the initial point and capture region.
2. Perform coarse tracking of the capture region.
3. Perform fine-tracking alignment.
4. Communicate the dynamic laser information between the two parties.

The keys to this process are the coarse and fine tracking. For coarse tracking, the scanning capture time is long and the capture probability is low, owing to the large uncertainty area and small beam dispersion angle. The solution is to use precision metering and pointing techniques to compensate for the capturing speed and vibration. For fine tracking, the tracking beam has a small beam dispersion angle and is easily affected by vibrations during the platform operation. The solution is to apply high frame rate camera detection to improve the communication effect and accuracy.

High-Rate, High Power-Emitting Lasers

To ensure that the system can effectively filter out background light, while providing the power required for the actual operation of space laser communication, space laser communication technology requires lasers with a high speed, high power, and narrow linewidth. Usually, high-speed laser emission technology involves external modulation, whereas high-power emission technology involves internal modulation.

For external modulation emission technology, the main technical difficulty is that the bias-voltage operating point of the electro-optical modulation crystal is greatly affected by temperature, that is, it has a high half-wave voltage at high power (inversely proportional to the length of the modulation crystal). The main technical approaches are to increase the peripheral temperature-controlled feedback circuit to suppress the bias-point drift, adopt a folded optical path to increase the optical range, reduce the half-wave voltage, and adopt multi-point lateral pumping to achieve high-power output.

For internal modulation emission technology, the main technical difficulty is meeting the high-speed, high-power, and high-quality requirements of the emitting laser source, that is, a low-power seed source, after amplification to meet the high

extinction ratio. The main technical approach is to develop a high-performance seed source laser and a high-gain amplifier.

Atmospheric Channel Impact Compensation Technology

It has been demonstrated through a large number of applications that laser beams in space laser communication will be affected by atmospheric turbulence when transmitting in the atmosphere. One example is laser energy attenuation, due to atmospheric absorption and scattering, which leads to system error codes; this is more obvious in high-speed laser communication. Another difficulty is atmospheric scattering and scintillation effects, which lead to laser spot-center shift and wavefront aberration problems.

To solve these problems, wide-range laser light-intensity stabilization and high-precision real-time wavefront aberration correction techniques can be used, respectively. The use of wide-range laser-intensity stabilization technology can ensure the signal-to-noise ratio of space laser communication. Using high-precision real-time wavefront aberration correction technology can suppress the influence of atmospheric turbulence on a transmitted beam wavefront. For example, multiaperture wavefront detection using a Hartmann sensor can correct wavefront aberrations to a certain extent.

However, the main difficulty of high-precision real-time wavefront aberration correction technology lies in the compensation of laser arrival-angle undulation, wavefront deformation, and the influence of the attached surface layer during air flight. These can be jointly corrected by introducing wavefront aberration compensation mirror technology into the detection system.

Summary

At present, radio communication is limited by frequency resources and transmission rates, and it has been difficult to meet the requirements of large-capacity data transmissions. Laser communication has a high transmission rate, strong anti-interference, and small size, which can break the low transmission rate bottleneck of traditional RF communication. In addition, unlike RF communication, laser communication is not regulated by the International Telecommunication Union (ITU) and does not require expensive band license fees, which gives it the advantage of a lower cost. Therefore, satellite laser communication has become an important means of high-speed communication for satellites and other space vehicles and has become a popular research topic in recent years.

This chapter compared the development of space laser communication technology at home and abroad and presented typical space laser communication demonstration tests. It analyzed the key technologies and technical difficulties, summarized the future development trends, and discussed directions for the future research and development of space laser communication.

Space laser communication is conducive to the realization of seamless global coverage, barrier-free communication, and the construction of an integrated global heaven-and-earth communication network. It also helps researchers in China understand the vision and planning of developed countries over time, which provides a reference for China's future space laser communication technology research.

References

1. Chen Zhihao. The invention of optical telephone, a milestone in the history of modern optical communication [J]. Optical Communication Technology, 1993, 17 (05):288–2810.
2. Wang Shiling, Fang Fengzhou. High-power lasers and their development [J]. Advances in laser and optoelectronics, 2017, 54(09):51–64.
3. Koepf G A, Marshalek R G, Begley D L. Space laser communications: a review of major programs in the United States [J]. International journal of electronics and communications, 2002, 56(4): 232–242.
4. Gerhard A. Koepf, Robert G. Marshalek, David L. Begley. Space Laser Communications: A Review of Major Programs in the United States [J]. AEUE - International Journal of Electronics and Communications, 2002, 56(4): 232–242.
5. Jeganathan M, Portillo A, Racho C S, et al. Lessons Learnt from the Optical Communications Demonstrator (OCD) [J]. Proceedings of SPIE - The International Society for Optical Engineering, 1999, 3615(1):23–30.
6. Barducci A, Casalbuoni R, Lusanna L.45-km horizontal path optical link demonstration [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2001, 9(4): 1102–1112.
7. Robert Pease. Optical laser-communications systems carve niche in metro markets [J]. Lightwave, 1999, 6(9):23–28.
8. Kim I I, Riley B, Wong N M, et al. Lessons learned for STRV-2 satellite-to-ground lasercom experiment [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2001, 4272:1–18.
9. Devoe C E, Pillsbury A D, Khatri F, et al. Optical overview and qualification of the LLCD space terminal [C]. International Conference on Space Optics 2016. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2018.
10. Caplan D O, Carney J J, Lafon R E, et al. Design of a 40-Watt 1.55μm uplink transmitter for Lunar Laser Communications [J] Proceedings of SPIE - The International Society for Optical Engineering, 2012, 8246: 82460M.
11. Boroson D M, Biswas A, Edwards B L. MLCD: overview of NASA's Mars laser communications demonstration system [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2004, 5339.
12. Boroson D M and Robinson B S. The lunar laser communication demonstration: NASA's first step toward very high data rate support of science and exploration missions [J], Space science reviews, 2014, 185(4):115–129.
13. Boroson D M, Robinson B S, Murphy D V, et al. Overview and results of the Lunar Laser Communication Demonstration [C]. Spie Lase. International Society for Optics and Photonics, 2016. 8971: 89710S.
14. Oaida B V, Wu W, Erkmen B I, et al. Optical link design and validation testing of the Optical Payload for Lasercomm Science (OPALS) system [J]. Proceedings of Spie the International Society for Optical Engineering, 2014, 8971.
15. Sputniknews. Us Marines Test Laser Communication System to Beat Radio Jammers[EB/OL]. https://sputniknews.com/military/201808241067456455-USMC-Tests-Laser-Communication-System/.

16. Google Alphabet successfully connects two countries across a river via an airborne beam up to 20Gbps+ [Z]. 2021: https://www.36kr.com/p/1402405703800196.
17. Biswas A, Srinivasan M, Piazzolla S, et al. Deep Space Optical Communications [C]//Free-space Laser Communication and Atmospheric Propagation Xxx: Proc. SPIE, 2019.
18. Edwards BL, Daddato R, Hamkins J, et al. An Update on the Ccsds Optical Communications Working Group [C], IEEE International Conference on Space Optical Systems and Applications, 2018.
19. Rose T S, Rowen D W, Lalumondiere S, et al. Optical communications downlink from a 1.5U Cubesat: OCSD program [C]. International Conference on Space Optics - ICSO.20110.
20. Toyoshima M, Yamakawa S, Yamawaki T, et al. Ground-to-satellite optical link tests between Japanese laser communications terminal and European geostationary satellite ARTEMIS [J]. Proc Spie, 2004, 5338:1–18.
21. Lange R, Smutny B, Wandernoth B. 142km, 5.625Gbps Free space optical link based on homodyne BPSK modulation [C]. Free-Space Laser Communication Technologies XVIII, Proceedings of SPIE, 2006, 6105: 61050A.
22. Gregory M, Hemmati H, Heine F, et al. Inter-satellite and satellite-ground laser communication links based on homodyne BPSK [J]. Proc Spie, 2010, 7587:75870E.
23. Hauschildt H, Mezzasoma S, Moeller H L, et al. European data relay system goes global [C]. 2017 IEEE International Conference on Space Optical Systems and Applications (ICSOS). IEEE, 2017:15–19.
24. Perdigues J M, Sodnik Z, Hauschildt H, et al. The ESA's Optical Ground Station for the EDRS-A LCT in-Orbit Test Campaign: Upgrades and Test Results [C]. Proceedings of SPIE, 2017, 10562:105622V.
25. Horwath J, Diaz Gonzalez D, Martin Navajas L, et al. Test results of error-free bidirectional 10Gbps link for air-to-ground optical communications [C]. SPIE, 2019.
26. Quintana C, Wang Q, Jakonis D, et al. A High Speed Retro-Reflective Free Space Optics Links With UAV [J]. Journal of Lightwave Technology. 2021, 39(18):5699–5705.
27. Komatu K, Kanda S, Hirako K, et al. Laser beam acquisition and tracking system for ETS-VI laser communication equipment (LCE) [J]. Proceedings of SPIE - The International Society for Optical Engineering, 1990, 1218:96–108.
28. Toyoshima M, Takahashi T, Suzuki K, et al. Laser beam propagation in ground-to-OICETS laser communication experiments [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2007, 23:65510A–65510A-13.
29. Yanko Noda, Shao Yi. Experimental research on optical communication on space station [J]. Advances in Laser and Optoelectronics, 2002, 39(1):15–110.
30. Perlot N, Knapek M, Giggenbach D, et al. Results of the optical downlink experiment KIODO from OICETS satellite to optical ground station Oberpfaffenhofen (OGS-OP) [C]. 2007: 645704-1–645704-9.
31. Toyoshima M, Shoji Y, Kuri T, et al. Development of Real-Time 3-Gbit/s BPSK Optical Coherent Receiver Using Field Programmable Gate Array for Free-Space Laser Communications [C]. 26th International Communications Satellite Systems Conference (ICSSC). 2009.
32. Teng Yunjie, Song Yangsong, Tong Shoufeng, et al. Capture performance study of laser communication system based on airship platform [J]. Journal of Optics, 2018, 38(06): 78–84.
33. Wu Feng, Yu Siyuan, Ma Zhongtian, et al. Correction of targeting angle deviation and in-orbit verification of satellite-ground laser communication link [J]. China Laser, 2014, 41(06): 154–1510.
34. Huang J, Deng K. Several theoretical issues in the research of atmospheric coherent laser communication [J]. Journal of Quantum Electronics, 2020, 37(05):556–565.
35. Nan Xu, Liren Liu, De'an Liu, et al. Optical phase-locked loops in interstellar coherent optical communication [J]. Advances in Laser and Optoelectronics, 2008, 45 (04):25–33.
36. Lin Y, Ai Y, Shan X. Identification of electro-optical tracking systems using genetic algorithms and nonlinear resistance torque [J]. Optical Engineering, 2017, 56(3):033105.

37. Shao-Wei Hao, Yong-Jun Li, Shang-Hong Zhao, et al. Optimal power allocation for inter-star visible light communication based on non-orthogonal multiple access [J]. China Laser, 2021, 48(07):125–134.
38. Shi Jizong, Zhu Ren, Zhao Siwei, Xiao Renliang, Zhong Chaoyang, Hou Xia, Chen Weibiao. Fast laser wavelength locking system based on piezoelectric ceramics and fiber grating [J]. China Laser, 2020, 47(12):16–21.
39. Mo Chen, Chao Liu, Daoman Rui, et al. Experimental results of 5-Gbps free-space coherent optical communications with adaptive optics [J]. Optics Communications, 2018, 418: 115–1110.
40. Zhang Jun-Xin, Mei Hai-Ping, Shen Liu-Jing, Huang Yin-Bo, Luo Fu, Wu Xiao-Yan. Statistical characterization of laser short-exposure spots in turbulent atmospheres in Gaomei Gu, Yunnan [J]. China Laser, 2021, 48(01):77–86.
41. Zu Jifeng, Liu Liren, Luan Zhu, et al. Advances and trends in interplanetary laser communication technology [J]. Advances in Laser and Optoelectronics, 2003, 40(3):7–11.
42. Zhang Jingang, Zhou Guangming, Zhang Daming, et al. Research and prospect of laser communication technology based on space-based platform [J]. Aerospace General Technology, 2022, 6(04):65–72.
43. Liu Liren. Satellite laser communication I link and terminal technology [J]. China Laser, 2007, 13(01):3–20.
44. Jiang Huilin, An Yan, Zhang Yalin, et al. Current status, development trend and key technology analysis of space laser communication [J]. Journal of Aircraft Measurement and Control, 2015, 34(03):207–218.
45. Ke Xizheng, Deng Lijun. Optical wireless communication [M]. Beijing:Science Press, 2016: 165–169.

Research Progress on Satellite Laser Communication Networks

2

Contents

Introduction 22
Satellite Laser Communication 22
Foreign Research Progress 23
Domestic Research Progress 26
Key Technologies for Satellite Laser Communication 29
  Key Optical Technology 30
  Key Signal-Processing Technologies 30
  APT Technology 30
  Opto-Mechanical Design Technology 31
Satellite Laser Communication Networks 31
  Satellite Orbits 31
  Satellite Network Architectures 32
  Satellite-Network Topologies 33
  Interstellar Laser Links 37
  Satellite-Routing Algorithms 39
Summary 41
References 41

Abstract

Satellite laser communication network through space laser communication technology make up for the use of microwave wireless communication in satellite network, while having the global coverage capability of satellite network; it has the advantages of high data transmission rate, large communication capacity, low power consumption, small antenna size, strong anti-interference ability, and so on. This chapter summarizes the research progress of satellite laser communication networks at home and abroad, introduces the key technologies of satellite laser communication, and describes important parts such as orbits, inter-satellite links, and routing algorithms of satellite laser communication networks. Finally, the full chapter is summarized and the development direction of satellite laser communication network is prospected.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_2

Keywords

Satellite laser communication · Satellite networking · Inter-satellite laser link

Introduction

Conventional satellites use Ka- and Ku-band microwaves for communication, which limits the communication capacity of communication satellites to a certain extent. Satellite laser communication has the advantages of being frequency independent, having a large bandwidth, high transmission rate, and good confidentiality; its load and communication terminals are both small in size, light in mass, and low in energy consumption [1].

Recently, there has been a global boom in satellite networking, and countries have invested in satellite laser communication network construction, competing for space spectra, orbits, and other resources. With the widespread application of multisource remote sensing technologies, such as high space, high spectrum, and high time domains, a large number of requirements for space-based information services have been proposed. However, it has been difficult for single-satellite relays to meet the growing service demand. Satellite laser networks, jointly constituted of multiple satellites, have become an inevitable trend for future communication development.

Satellite Laser Communication

Satellite laser communication uses a laser beam as a carrier wave for direct two-way transmissions of voice, data, and image information in space. Laser communications can be divided into two categories: space and atmospheric. The former is mostly used for inter-satellite communication, interstellar communication, spacecraft, and space stations in vacuum conditions; the latter is used in a wide range of applications, mainly for connections between satellites, land and sea users, and aircraft.

A satellite laser communication system consists of three parts: the ground station, satellite, and user terminal. The ground base station uses an optical antenna to send a laser signal to the satellite, which receives the signal and sends it to the user terminal. They are widely used in inter-satellite communication, remote sensing, and navigation.

Compared with traditional microwave communication technology, satellite laser communication technology has the following advantages:

1. Large communication capacity: The frequency of a laser is three to four orders of magnitude higher than that of a microwave, and the frequency band is wider, allowing for large-scale data transmissions in a shorter period.

2. High communication rate and low power consumption: High-speed communications of 10 Gbit/s or more can be achieved. Owing to its highly concentrated energy and small beam-divergence angle during transmission, a laser does not diffuse easily and consumes less energy than microwave communication.
3. Strong resistance to interference: The laser beam-scattering angle is very small, which is not easy to detect or interfere with.

Satellite laser communication technology also has the following shortcomings and problems:

1. Vulnerability to external environmental factors. These include the influence of the atmospheric environment, space ambient-temperature changes, background light, platform vibration, and relative motion.
2. Sensitivity to external environments, such as atmospheric environments: Fog, rain, snow, and hail can attenuate a laser transmission in the air; thermal coronas, atmospheric turbulence, and airborne suspended matter can all reduce the quality of a laser transmission.
3. Influenced by temperature change: Because each functional module in the laser communication system is located at different locations of the load terminal, a temperature gradient will be generated. Each module material is different, and the coefficient of thermal expansion is also different. A temperature gradient will lead to the deformation of the mechanical structure, which has a greater impact on the installation, as well as the acquisition, pointing, and tracking accuracy of the system.
4. Affected by background light: Background light, such as daylight and moonlight, may interfere with interstellar links, thereby reducing the transmission quality and reliability of the communication system.

Foreign Research Progress

In 1995, the theoretical concept definition and transmission layer design of satellite laser communication networks were investigated by the Information and Decision Systems Laboratory at MIT, and Chan et al. completed a large number of theoretical design definitions and other work related to satellite laser networks, which were important for the development of satellite laser communication networks [2].

In 2000, Karafolas et al. proposed an optical wavelength division multiplexing (WDM) network for application to satellite constellations, analyzed the constellation characteristics of interstellar laser links, and presented the existing problems [3].

In 2001, Ors et al. investigated methods to apply Internet-protocol quality of service (IPQoS) to satellite networks using multiprotocol label switching and described the problems and solutions of this technique [4].

In 2001, Ekici et al. proposed a new distributed routing algorithm called the datagram routing algorithm. This routing algorithm combines virtual nodes with distributed policies to calculate the shortest delay path; it only needs to know the location of the destination satellite to transmit the packet directly [5].

In 2004, Donner et al. proposed a satellite networking concept based on multi-protocol tag switching and explored various networking schemes, according to the characteristics of a dynamic topology [6].

In 2007, Papapetrou et al. proposed a location-assisted distributed routing algorithm that utilized the concept of wireless self-organizing networks to balance the transmission load and increase the data transmission rate [7].

In 2007, Papapetrou et al. applied a location-assisted on-demand routing algorithm to satellite communications. Simulation results showed that the algorithm could reduce end-to-end delay and delay jitter, as well as improve the transmission efficiency of the system. They also designed a deterministic satellite-movement mechanism, based on the number of satellites involved in a traffic transmission [8].

In 2007, Korçak et al. proposed a new priority-based adaptive shortest-path routing algorithm that ranked the historical utilization of links and buffer information to select the appropriate next-hop link from multiple shortest links in a distributed manner. It effectively solved the network-congestion problem in high-latitude regions [9].

In 2011, Fu et al. proposed a routing algorithm for traffic engineering in multi-protocol label-switched networks, which could dynamically establish interstellar laser links with bandwidth guarantees to balance the network load [10].

In 2012, Kawamoto proposed a method for solving interstellar laser link blocking by recalculating the routing table at the ground receiver. This method improved the link-usage efficiency; however, it also introduced a large signaling overhead [11].

In 2013, Pan et al. proposed a minimum-contention routing algorithm that increased the utilization of network resources to a new level and allowed the network to accept more routing requests [12].

In 2014, Chaudary et al. investigated free-space optical communication based on orthogonal amplitude modulation (OAM) and orthogonal frequency-division multiplexing. They used a 10-Gbit/s data stream with a four-level OAM sequence transmitted through a free-space optical communication system in different atmospheric environments. The results showed that the maximum transmission distance was 185 km under ideal weather conditions and 2.5 km in fog [13].

In 2016, Fernls R et al. described the benefits of applying software-defined network/network-function virtualization techniques in satellite laser communications, proposed and analyzed three scenarios, and briefly described the challenges faced by this technique [14].

In 2016, Huang et al. proposed a virtual topology-based routing algorithm to address the challenges posed by the dynamic topology of low earth-orbit (LEO) satellite networks, which discretized the dynamic topology of the satellite network in time, where the topology of the satellite network was constant in each snapshot. The signaling gateway station could calculate the routing table in each snapshot in advance, without the need for real-time satellite calculations [15].

In 2017, Gardisis et al. investigated the application of software-defined networking/network-function virtualization techniques to the satellite communication domain, analyzing the benefits and challenges of integrating satellite infrastructure into software-defined networks [16].

In 2017, Lyras et al. developed a model for cloud-free line-of-sight probability prediction for single optical satellite links and multiple spatially separated optical satellite links by considering a combination of three perspectives: elevation angle, ground-station altitude, and cloud spatial variability [17].

In 2017, Khallaf et al. proposed a modulation method to create a hybrid of orthogonal amplitude modulation and multipulse position modulation over turbulence-free and gamma–gamma free-space optical channels, which had an improved bit-error rate (BER) performance compared to previous modulation methods [18].

In 2018, Lyras et al. analyzed the variation patterns of cloud-coverage statistics, which provided a theoretical basis for selecting optical ground terminals [19].

In 2018, Liu et al. investigated the joint layout of controllers and gateways in a converged network of 5G terrestrial and satellite networks, based on software-defined networks, and provided a solution that could obtain near-optimal delay, as well as maximum network reliability while satisfying delay constraints [20].

In 2019, Wang et al. proposed an adaptive routing algorithm that was applicable to a three-layer satellite architecture. It optimized information links with dynamic changes in the satellite topology to improve the dynamics and flexibility of satellite networks [21].

In 2019, Yang et al. proposed and analyzed a K-means network-partitioning algorithm for integrated satellite–ground networks combined with software-defined networks, which improved the network reliability performance and reduced transmission delay [22].

In 2019, Lyras et al. provided two optimization algorithms for medium earth-orbit (MEO) satellite networks operating in the optical band, both of which ensured the effectiveness of interplanetary links in the presence of cloud blocking and minimized the network size of optical ground stations [23].

In 2019, Liu et al. proposed a routing algorithm based on segmented routing for traffic return to address the load imbalance in LEO satellite networks caused by the uneven distribution of users in various regions of the world. Specifically, heavy- and light-load areas in a recursive network were dynamically divided according to the relative position relationship between the gateway and the reverse seam, and the corresponding routing strategies were adopted in different areas [24].

In 2020, Liu proposed a strategy based on a selective iterative Dijkstra algorithm and selective split-load balancing, which fully exploited the inter-satellite laser links in LEO satellite networks. In addition, for the link-congestion problem caused by uneven traffic distribution in LEO networks, this strategy could reasonably divert the congested nodes' traffic and reduce the load on other regional networks [25].

In 2021, Polnik et al. proposed two types of uncertainty models, based on the predicted cloud volume: robust optimization of polyhedral uncertain sets and robust optimization of matrix-based fuzzy sets. The performances of the different models

were analyzed and calculated, and based on this, a 6-month UK work plan for artificial satellites and ground-based observatories was developed [26].

Domestic Research Progress

In 2004, Chan et al. studied the constellation topology of space-based backbone networks and proposed a fast air-to-air communication scheme for handling space-based resources in orbit [27].

In 2006, Li et al. proposed a method to realize inter-satellite laser link communication networking based on WDM technology and established a WDM/optical code-division multiple-access-based satellite network by analyzing the topology of a satellite laser communication system combined with optical code-division multiple access [28].

In 2007, Yi et al. proposed a routing algorithm for a geosynchronous earth orbit (GEO)/MEO two-layer satellite network architecture, which uploaded data information from the MEO layer to GEO satellites through cluster management and cluster-head backup. The GEO satellites propagated MEO-satellite information to each other to better manage the lower-layer satellites [29].

In 2008, Qi et al. proposed a routing algorithm based on offline computation for LEO constellations. It provided a traffic-adaptive mechanism using backup paths while ensuring routing efficiency. Simulation results showed that the algorithm required only small on-star storage and processing overhead and had good end-to-end delay performance [30].

In 2009, Yang et al. studied satellite laser networks with a WDM architecture and found that network crosstalk had a significant impact on the network performance; simulation results showed that increasing the laser bandwidth effectively reduced the crosstalk [31].

In 2009, Li et al. proposed a global continuous-coverage LEO/MEO double-layer satellite optical network structure using the coverband method. The simulation results showed that the coverage of orbiting satellites in this structure reached 910.9% globally and 100% for domestic ranges [32].

In 2010, after analyzing the wavelength demand for laser transmissions in a non-geosynchronous satellite network, Tan et al. identified the factors affecting the wavelength demand. Simulation results showed that the wavelength demand was related to the percentage of active terminals, the maximum number of allowed inter-satellite link hops, and the inter-satellite link terminals per satellite [33].

In 2010, Tang et al. investigated satellite laser networks using wavelength routing, proposed an algorithm that could label their wavelength demand, and estimated the boundaries of the wavelength demand, based on the proposed wavelength-estimation algorithm [34].

In 2010, Li et al. proposed a set of two-layer LEO/MEO satellite laser networks with global coverage based on a zero phase factor. The network provided minimum pitch angles of 21° and 16° for single-satellite and double-satellite systems in China, respectively. Simulation results showed that the LEO constellation with a zero phase

factor had a great advantage in several performance parameters, such as azimuth and pitch angle. The difference between short- and long-wave bands was small at low traffic, which is well suited for inter-satellite laser links [35].

In 2014, Gao et al. proposed a routing algorithm that changed the manner in which the routing tables were calculated. Simulation results showed that the algorithm was only suitable for satellite networks with a small number of satellites and a simple topology [36].

In 2014, Yang et al. proposed a bandwidth-estimation transmission-control protocol algorithm based on link length. By optimizing the relevant parameters in the congestion window, slow-start threshold, and time-out retransmission timer, the congestion-control algorithm ensured high throughput during abrupt changes in transmission delay and was more suitable for the dynamic environment of satellite networks. Simulation results showed that the throughput performance of the bandwidth-estimation transmission-control protocol algorithm phase in the satellite network was improved by 2.8% compared to that of the transmission-control protocol algorithm, which effectively solved the problems caused by sudden changes in delay [37].

In 2015, Dong et al. investigated an ant-colony algorithm based on a small-window strategy to solve the problems of routing and wavelength assignment in satellite laser networks. Compared with Dijkstra's algorithm, the ant-colony algorithm reduced the blocking probability by up to 50% and improved resource utilization, especially in the case of low wavelength utilization, by up to 45% [38].

In 2015, Chen et al. designed a 2+2 satellite networking architecture, which was combined with the analytical approach of terrestrial-programmed switching technology to define the performance metrics of optical networks. Simulation results showed that this satellite network could support random access to 50 user satellites, while satisfying certain access-loss probabilities and resource utilization [39].

In 2015, Chen proposed a two-layer satellite-network routing algorithm using the predictability of the satellite position, which solved the problem of data loss due to inter-satellite links being cut off when in-orbit satellites entered higher-latitude regions. This greatly reduced the packet-loss probability of the laser network while improving the system reliability [40].

In 2017, Zheng et al. studied the topology of heterogeneous satellite laser networks, addressing the initialization and reconfiguration of the relaxation solution [41]. They also designed a distributed minimum spanning-tree algorithm for a distributed satellite laser network that improved the overall network performance and reduced the computational complexity of the system [42].

In 2017, Wen et al. proposed an ant-colony routing and wavelength assignment algorithm to address the negative effects caused by Doppler-wavelength shifts in satellite laser networks. Simulation results showed that the algorithm performed well in terms of communication success probability and had a good global-search capability [43].

In 2017, Xu et al. proposed an intelligent capacity-analysis method for satellite networks, designed a set of capacity models for satellite communication systems based on their characteristics, and introduced them to the capacity analysis of

satellite communication systems to achieve high computational efficiency and reduce the time overhead [44].

In 2018, Li et al. studied an optical burst-switching architecture for GEO/LEO two-layer satellite networks, focusing on the design of a resource-scheduling mechanism for space-based backbone networks, based on optical burst switching, to provide high-speed data transmission rates and high switching capacity [45].

In 2018, Liu et al. proposed a swarm-optimization algorithm to reduce the cost of laser links and comprehensively consider wavelength-resource utilization and communication hops to efficiently utilize wavelengths, reduce unnecessary bypassing, and guarantee the BER performance of the system [46].

In 2018, Liu et al. established a satellite laser network model and proposed an algorithm based on subtree merging and maximum algebraic connectivity. It effectively improved the algebraic connectivity of the satellite laser network while ensuring that the network topology had higher average edge weights [47].

In 2018, Zhang et al. proposed a multibeam switching mechanism to solve the problem of frequent beam switching due to software-defined networks, based on a traditional satellite network switching mechanism. It achieved high-precision and fine-grained switching by quantitatively analyzing the multiple metrics involved in the switching process [48].

In 2020, Liu et al. proposed a dynamic laser network wavelength-estimation method for the wavelength multiplexing problem in intermittent interstellar links and further analyzed the impact of topology changes on the wavelength demand. The simulation results showed that the higher the network connectivity, the better the network reachability. The wavelength requirement depends on the network connectivity and required reachability time [49].

In 2020, Peng et al. proposed a split-traffic sparse method, which is an algorithm that can satisfy the service traffic demand as much as possible while minimizing the number of wavelengths. It is suitable for sparing static traffic in large networks. This provides an effective approach for the sparse traffic design of wireless networks [50].

In 2020, Sun et al. proposed a super-satellite node computational storage-function algorithm that effectively reduced wavelength requirements and improved network connectivity. This study reduced the cost of onboard lasers, modulators, and establishing inter-satellite links, which was an important reference for establishing satellite laser networks [51].

In 2021, Zhang proposed two satellite communication-network routing algorithms around a low-orbit satellite network and low- and medium-orbit satellite network scenarios to address network-node failures and network-congestion problems [52].

In 2021, Wang et al. proposed a method to improve the capacity of satellite networks using network coding and divided it into two cases: network coding and non-network coding. Simulation results showed that the improvement in network capacity with the network-coding scheme significantly exceeded that with the

non-network coding scheme, and the improvement in network capacity increased with an increase in resources and network size [53].

In 2022, Liu et al. proposed a beacon step-tracking algorithm. Simulation results showed that the proposed tracking method was closer to the tracking performance of the program when the carrier's position and attitude changes were small, and it reduced the service-interruption probability when the user terminal accessed the satellite communication network [54].

In 2022, Zuo et al. proposed a path-selection method for dynamic LEO satellites, which used a path-selection model, built using deep learning, to automatically select paths based on information, such as spatial location, spacing, and available bandwidth of peripheral nodes. Simulation results showed that the algorithm had better delay characteristics, convergence, and generalization capabilities than typical distributed routing algorithms [55].

In 2022, an effective load-balancing algorithm was proposed by Zheng et al. for satellite–ground networks to reduce the propagation delay caused by high loads on ground network links. Simulation results showed that setting an adaptive overload threshold could effectively reduce the link congestion [56].

In 2023, Liu et al. proposed a star cross-cycle link-building strategy (SCLBS) to solve the reverse seam link-building problem for polar-orbiting constellations. Simulation results showed that SCLBS could reduce the data transmission delay of polar-orbiting constellations by 30–60%, compared with previous link-building strategies, which significantly improved the transmission performance of LEO satellite networks [57].

In 2023, Xie et al. proposed a quantum evolutionary algorithm based on an excitation function and processing delay to solve the distributed inference problem of deep neural networks in LEO satellite scenarios. Simulation results showed that the algorithm had both lower delay characteristics and higher overall utility, which could improve the efficiency of intelligent services in low-orbit satellite scenarios [58].

Key Technologies for Satellite Laser Communication

Satellite laser communication is a complex and difficult technology that involves many fields and disciplines, such as optics, mechanics, signal processing, mathematics, and computers. Optical technology mainly refers to high-power light-source technology and the design of high-quality optical systems. Signal-processing technology mainly includes modulation and demodulation, suppression of background noise, compensation for atmospheric-channel effects, and high-sensitivity detection. Mechanical aspects are mainly for the compensation of platform vibration and attitude and high-precision acquisition, pointing, and tracking technology (APT). This section focuses on key technologies in satellite laser communication research.

Key Optical Technology

Optical systems mainly involve wavelength selection and optical-path design. When selecting a wavelength, it is necessary to reduce the loss of optical signals through the atmosphere and the influence of solar radiation, as well as ensure a high response to the detector. With the continuous development of communication services, satellite optical-communication systems often require multiple wavelengths to be transmitted in parallel to guarantee service-transmission quality and bandwidth.

The wavelengths often used between the satellite and ground link are 1550 nm, 850 nm, and 1000 nm. Specific wavelengths should be selected and designed according to the different transmission methods, transmission environments, and service requirements. In satellite laser communication systems, optical antennas are used to transmit and receive optical signals. Traditional optical-antenna technology has become more mature; however, with the development of miniaturized laser communication satellites, the volume and weight of traditional satellites cannot meet these new requirements.

Key Signal-Processing Technologies

The choice of optical-signal modulation and demodulation technology in satellite laser communication requires the consideration of several aspects, including optical-power efficiency, frequency-band utilization, information transmission rate, and interference resistance. Spatial filtering and signal modulation methods are commonly used to reduce the impact of background noise on optical-signal transmissions. Factors, such as the angle of arrival, Doppler shift, and laser linewidth, were considered in the filter design [59].

One of the most widely used modulation techniques for suppressing background noise is multipulse position modulation, which is often used in conjunction with other modulation methods to improve the transmission efficiency of optical signals.

APT Technology

Highly efficient acquisition, pointing, and tracking technology is key to realizing satellite laser communication; it is especially the top priority when building a satellite–ground link. The system consists of two parts: a coarse-tracking system in the acquisition phase and a fine-tracking system in the pointing and tracking phases. The coarse-tracking system is first used to scan a large area, and when the transmitted signal is captured, the fine-tracking system is used to scan a small area. Using these two tracking systems in combination, fast and robust signals are captured.

Opto-Mechanical Design Technology

To reduce the power loss of a laser transmission in free space and increase the gain of the optical system at the transmitting end, the transmitted beam must be emitted near the diffraction-limit angle. Higher design requirements for fiber coupling, beam shaping, and telescope-face shapes are required to obtain a higher transmit-beam gain without affecting the transmission aperture.

When different core diameters, fiber beam-dispersion angles, and optical-system matching-selection methods are used for further research, near-diffraction-limit angle emission technology, launch-end laser-shaping collimation technology, and high-efficiency fiber-coupling technology can achieve breakthroughs. In terms of optical-infrastructure platform technology, a modular and lightweight optical system design is needed, which can meet the need for future simultaneous one-to-many transmissions in space laser communication networks [60].

Satellite Laser Communication Networks

Satellite Orbits

Satellite orbits can be classified according to their orbital altitude and provide the following services.

LEO: Low earth orbit. This is an Earth-centered orbit with an altitude of no more than 2000 km between the earth and the inner Van Allen radiation belts. LEO requires the lowest amount of energy for satellite placement. It provides high bandwidth and low communication latency. The satellites and space stations in LEO are more accessible to astronauts and service personnel.

LEO satellites have a small instantaneous field of view and can only observe and communicate with a small portion of the earth at a time, which means a "constellation" is needed to provide continuous coverage. LEO satellites also have a rapid orbital decay rate and require periodic orbital boosts to maintain a stable orbit or to replace old satellites as they re-enter the atmosphere.

MEO: Medium earth orbit. The orbit altitude is located between LEO and GEO, with a typical orbit altitude of 2000–36,000 km. Compared with LEO, it requires fewer satellites for global coverage and is therefore easier to manage, without the frequent-turnover problem of LEO satellites. Compared with GEO, the number of satellites required to provide global coverage is slightly higher because of the lower orbit altitude and the poorer coverage capability of a single satellite. However, the launch cost of a single satellite is lower, and it has stronger coverage capability in the bipolar region.

GEO: Geosynchronous orbit. This orbit has a constant orbital altitude, that is, the distance between the orbit and the surface is 35,786 km, and all GEOs share the same semi-long axis length. The orbital period of all GEOs is exactly equal to one sidereal day, which means that the satellites in this orbit return to the same point above the earth's surface every day, regardless of other orbital characteristics. Satellites in this

orbit have the largest single-coverage area compared to satellites in other orbits; however, the orbit is farther from the ground, resulting in the largest signal delay.

Satellite Network Architectures

The most widely used satellite network architectures are single layer and multilayer. At this stage, the research on single-layer satellite networks is more mature. However, with the rapid development of satellite laser communication network technology, single-layer satellite networks can barely handle the increasing service demand, and multilayer satellite networks are gradually becoming an important research direction for satellite laser communication networks at present.

Single-Layer Satellite Networks

A single-layer satellite network is a network of satellites that are in the same orbit. These are primarily the LEO, MEO, and GEO networks.

LEO satellite networks generally consist of hundreds of satellites. Because of their low orbit, the satellite–ground link has better performance and lower transmission delay. However, LEO satellite networks require a long time to build, and the large number of satellites leads to a complex satellite network-control system with a high cost, frequent satellite-network switching, and a short beam-switching interval. Because of the large number of signal gateway stations, the LEO satellite network has high requirements for APT systems, and the impact of the Doppler effect must be considered.

MEO satellite networks generally consist of a dozen satellites. Considering the inter-satellite link distance, uplink, and other factors, MEO satellite networks have a much higher transmission delay compared to LEO satellite networks; however, the switching probability is lower, the Doppler effect is reduced, and the space-control and APT systems are simpler [61].

The GEO satellite network has a high global coverage with a small number of satellites. The low switching probability and simplicity of the APT system are also advantageous. Compared with the MEO and LEO satellite networks, the laser link has the farthest transmission distance, resulting in a large link loss and transmission delay, which can only meet the needs of low-orbit high-power services. The coverage of middle- and high-latitude regions is not ideal.

Multilayer Satellite Networks

Because of their limitations, it is difficult for single-layer satellite networks to meet the needs of the current dramatic increases in service volume, higher transmission quality, and global coverage. Multilayer satellite networks make full use of the advantages of each layer of the satellite network and have better performance than single-layer satellite networks in several aspects, such as spectrum utilization, coverage elevation angle, and weight number [62]. The main multilayer satellite networks are double-layer and three-layer types.

Double-layer satellite networks are mostly LEO–MEO. The MEO satellites are connected to each other by an intra-layer inter-satellite link and to the LEO satellites within their visual range by an in-orbit inter-satellite link; the LEO satellites are also connected to each other by an inter-satellite link. LEO layer satellites are responsible for communicating with small mobile terminals to reduce link losses, whereas MEO layer satellites act as relays and provide communication between ground stations and large terminals [63].

A three-layer satellite network combines the advantages and shortcomings of both LEO and GEO network layers. The parent satellite manages the three-layer satellite network and the subsatellite handles the service requirements, thereby solving the problems of signal fading, time delay, and frequent switching. However, a three-layer satellite optical network system is large-scale and complex in structure and contains many satellite nodes, which increases the calculation volume and complexity of the system routing table. In addition, to reflect the routing status of the three-layer satellite optical network in real time, the routing table must be updated frequently, and the repeated calculation of the routing table leads to the generation of new link delays. The accumulation of new delays creates an additional load on the satellite network, increases communication overhead, and can even lead to link blocking.

Satellite-Network Topologies

Network Topology

The topology of a network refers to the interconnection relationship between the nodes in the network, which determines the number of communication links between each node, the transmission path, and the distance traveled by data and messages. Analyzing the network topology is the first step in designing the network, which is directly related to the overall network performance and is also the basis for implementing multiple communication protocols.

The most common satellite-network configurations are star, ring, and mesh topologies. More complex satellite networks are typically formed by combining these three basic network topologies. The following section discusses these three topologies.

Star Topology

In a star topology, all slave nodes in a network are connected to a central node located at the center of the network; however, the slave nodes are not connected to each other. The information transmission between the slave nodes depends on the central node, as shown in Fig. 2.1.

The star topology is more suitable for certain application-oriented satellite networks, which generally have one satellite as the core; the other satellites communicate with each other by relaying through the central satellite. In the star network, satellites in geosynchronous orbits mainly predominate, but satellites may also exist in other orbits.

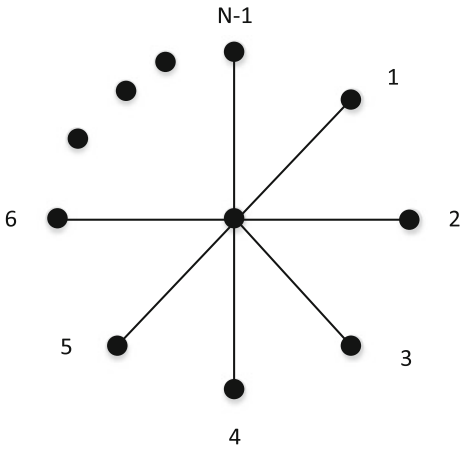


Fig. 2.1 Star topology [64]

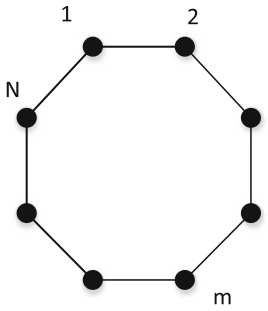


Fig. 2.2 Ring topology [64]

Ring Topology

A ring topology is a static interconnected topology consisting of multiple nodes interconnected in a certain order in a point-to-point manner to form a ring. Ring topologies can be classified as single ring, double ring, modified double ring, banded string ring, etc.

As shown in Fig. 2.2, the single-ring structure is highly symmetric and consists of N nodes, each of which is connected to a neighboring node through two ports, forming a closed-ring system. In an undirected ring, the vertex degree of each node is two, and the diameter of the network is $[N/2]$. In contrast, for a single-loop network, the diameter of the network increases linearly as the number of nodes N increases. In addition, because there are no redundant paths in a single-loop network, the entire system does not function properly if an error occurs in one link.

Figure 2.3 shows a band-chordal ring topology, i.e., a topology consisting of two nodes separated by a certain distance in a single-ring topology. The band-chordal ring topology is simple and reduces the diameter of the linear array and ring network.

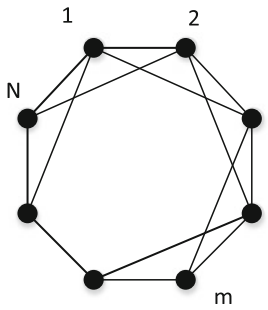


Fig. 2.3 Topology with chordal ring [64]

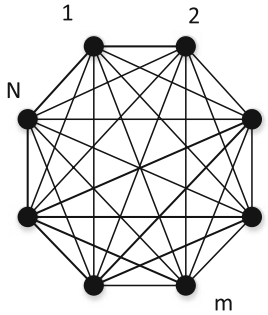


Fig. 2.4 Fully connected topology [64]

As shown in Fig. 2.4, for the banded chordal-ring mesh structure, the vertex degree increases and the mesh hole decreases. When the vertex degree increases to $N-1$, the network diameter reaches a minimum value of 1, the number of links is $N(N-1)/2$, and all nodes in the network are interconnected; this network structure is called a fully connected network.

In theory, a fully connected network is the ideal network topology; that is, every node in the network is connected to the others. Using this method, information can be transmitted between any nodes without passing through intermediate nodes. It can minimize communication delays and maximize point and edge connectivity; this means that fully connected networks have the strongest resistance to destruction, compared with other network structures.

However, in practice, these interconnections are not practical. When all nodes in the network are connected to each other, the number of node ports grows linearly with the number of nodes, but the number of links grows squarely. This makes the entire network more complex, difficult to implement, and expensive.

The advantage of the ring topology is that the structure, routing method, communication interface, and network management are relatively simple and easy to implement. In addition, multiple ring–ring constellations are interconnected, which can achieve a larger range of ground coverage. At present, the ring topology is commonly used in medium- and high-orbit satellite networks to cover the ground

using the minimum number; however, there are problems, such as low transmission bandwidth and large delays.

The shortcomings of the ring topology are that when the number of nodes in the network is too large, the system transmission efficiency decreases and the network response time increases. In addition, if a new satellite joins, the loop needs to be cut off, which is unfavorable to expanding the system.

Mesh Topology

A mesh network topology is composed of neighboring nodes connected to each other in a planar shape. The mesh topology is also called a distributed network topology because each node can connect to other nodes in the network; thus, there are sufficient redundant paths to each node to enable perfect reliability and high fault tolerance.

Various mesh topologies exist, primarily quadrilateral, hexagonal, and triangular, as shown in Fig. 2.5. To reduce the diameter of the network, a new topology can be formed by connecting the nodes located at the periphery of the network in a specific manner, such as a barrel network (Fig. 2.6), wreath network (Fig. 2.7), closed-spiral network, or twisted-spiral network.

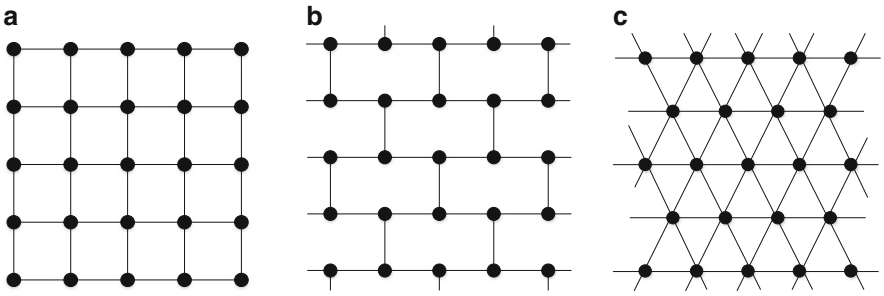


Fig. 2.5 Mesh topology [64]. (**a**) Quadrilateral grid; (**b**) Hexagonal grid; (**c**) Triangular grid

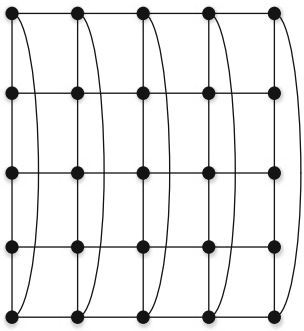


Fig. 2.6 Barrel network topology [64]

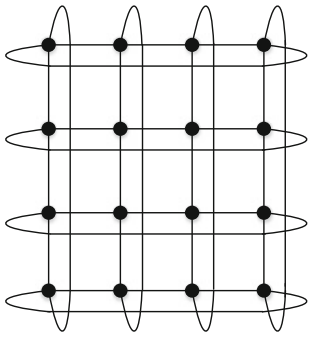


Fig. 2.7 Wreath network topology [64]

The mesh topology has sufficient redundancy, high system reliability, and better scalability with interstellar links. It has the advantages of a higher transmission bandwidth, higher data transmission rate, and lower latency and can achieve world-wide coverage. A mesh topology is typically used to achieve global communication and network support. However, a major drawback of this mesh structure is that it is expensive to build and requires multiple satellites.

Interstellar Laser Links

Interstellar Laser Link Geometry

Multiple satellite nodes are connected using inter-satellite links to form a satellite network. If the number of in-orbit satellites and inter-satellite links is the same, a satellite network with a regular mesh structure can be formed using inter- and intra-orbit links. Satellites in the same orbit are connected by intra-orbit links, which can be maintained durably because each satellite in a circular orbit is in a fixed position. The links between the orbits are not permanent.

When approaching high latitudes, the relationship between neighboring satellites changes, requiring on-board devices capable of turning 180° to track neighboring satellites and maintain the inter-orbit links. It is also possible to cut the original link and rebuild the link when the neighboring nodes change. As the distance decreases, the relative speed of the two satellites increases, and the Doppler shift increases.

Three inter-satellite link geometries are shown in Fig. 2.8: Fig. 2.8a shows a "W" geometry with four inter-satellite links; Fig. 2.8b shows a "tilted" geometry with four inter-satellite links per satellite; in Fig. 2.8c, each satellite has three interstellar links.

Interstellar Laser Link Performance Calculation

The transmitting and receiving gains of the antenna are G_T and G_R, respectively, and the optical transmitter power is P_T; thus, the laser power P_R received by the receiver antenna can be calculated [65]:

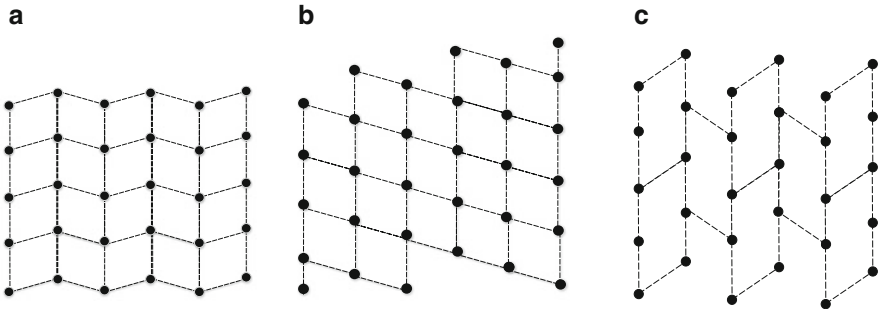


Fig. 2.8 Schematic diagram of interstellar-link geometries [64]. (**a**) W geometry; (**b**) Tilted geometry; (**c**) Three interstellar links

$$P_R = P_T + G_T + G_R - \left[L_P + \sum\nolimits_L + S_F\right] \tag{2.1}$$

where the standard deviation $S_F = 5$ dB, and L_P is the path loss:

$$L_P = \left[82 + 20\lg\frac{d_c}{\lambda}\right] dB \tag{2.2}$$

where d_c is the interstellar distance in kilometers (km) and λ denotes the laser wavelength in meters (m).

Σ_L includes other losses in addition to path loss. Because the position of the satellite is outside the atmosphere, there is no need to consider the attenuation, refraction, bending, or beam tearing of an atmospheric laser during the transmission of a light wave over a laser link. Only the absorption of light by optical elements need be considered. It can avoid beam-pointing errors due to the instability of the satellite attitude; this loss has no fixed value to use. It can be calculated by referring to the equation in [65]:

$$\begin{aligned}\sum\nolimits_L &= L_{TOA} + L_{ROA} + L_{DF} + L_{PT} + L_M = 1.5 + 1.5 + 1 + 1.5 + 1.5 \\ &= 7(dB)\end{aligned} \tag{2.3}$$

L_{TOA} and L_{ROA} are, respectively, the transmitting and receiving optical absorption losses of the transceiver antenna, L_{DF} is the scattering loss, L_{PT} is the pointing and tracking loss, and L_M is the outlier mixing loss.

In Eq. 2.1, if the antenna is a lens or reflection type, both G_T and G_R can be calculated using the following equation [65]:

$$G_i = 10\lg\left(\frac{\pi D_i}{\lambda_i}\right) \tag{2.4}$$

where D_i is the antenna aperture in meters (m).

A PIN photodetector is used for outlier-synchronization detection on an interstellar laser link with a signal to noise ratio S/N of [65]

$$\frac{S}{N} = \frac{\xi P_R}{\Delta f h_\nu} \tag{2.5}$$

where the quantum efficiency ξ is the number of electrons that can be produced per photon at the PIN; ν is the incident-light frequency expressed in hertz (Hz); $h = 6.626 \times 10^{-34}$ J·s, which is Planck's constant; Δf is a one-sided bandwidth, and the baseband-signal code rate $r_b = 2\Delta f$.

The BER P_e can be expressed as [65]

$$\begin{cases} P_e = \frac{1}{2}\, erfc\left(\frac{Q}{\sqrt{2}}\right) = \frac{1}{2}\, erfc(Q') \\ Q' = \frac{Q}{\sqrt{2}} \end{cases} \tag{2.6}$$

where Q is the value proportional to S/N, and *erfc* is the complementary error function.

If P_e is set to 10^{-9}, $Q' = 6$, in connection with Eq. 2.5; setting $\xi = 1$, the required P_R is

$$P_R = \left(\frac{S}{N}\right) \cdot h\nu \cdot \Delta f / \eta = 18 h\nu \cdot r_b \tag{2.7}$$

The interstellar transmission rate of the corresponding interstellar laser link can be calculated as follows:

$$r_b = \frac{P_R}{18 h\nu}. \tag{2.8}$$

Satellite-Routing Algorithms

According to the different aspects of routing decisions, satellite-routing algorithms can be classified as centralized, distributed, or hybrid. Centralized routing means that a central node in the network calculates the routing table and distributes it to each satellite node. Distributed routing means that each satellite node is the subject of the routing decision and can independently calculate the best node for the next hop and send the data. Hybrid routing combines the centralized calculation of centralized routing and the independent decision feature of distributed routing.

Centralized Satellite-Routing Algorithm

The centralized routing algorithm assigns all the calculation and maintenance of the routing table to the central node and distributes it to each satellite node. The satellite

node selects the best node for the next hop against the routing table and transmits the data. The central node collects the information of inter-satellite links with the dynamic change of the network-topology mechanism, completes the calculation of the routing table, and updates it [66].

From a global perspective, a centralized routing algorithm can fully consider the overall network state and routing-table calculation of the satellite network, thus achieving better traffic distribution and load balancing. However, the limited on-satellite processing and computing capability of a single satellite causes the centralized routing algorithm to suffer from high computational complexity and high signaling overhead. Even if a ground station with strong computing power is used as the central node, it is difficult to monitor the state changes of links and nodes and update the routing tables in real time, when satellite nodes or links are abnormal or fail. This has an unpredictable impact on the network performance.

Distributed Satellite-Routing Algorithm

The difference between distributed routing algorithms and centralized routing algorithms is that in the former, each satellite node makes its own routing decisions and can independently calculate the best node for the next hop and send data. Consequently, the routing table can be calculated without going through the central node [67].

The distributed routing algorithm considers the problem of low processing and computing capacity on a single satellite, distributes the computing volume of a single satellite to multiple satellites, reduces the computing burden on the satellite nodes, enables better monitoring of network-traffic changes, and dynamically updates the routing tables.

Distributed routing algorithms also have many shortcomings. Because the satellite nodes select the next hop independently, the path chosen by the routing decision is not necessarily the optimal solution from an overall perspective. This may lead to data being forwarded to nonoptimal paths and increase unnecessary signaling overhead.

Hybrid Satellite Routing Algorithm

The hybrid routing algorithm combines the advantages of centralized routing and distributed routing. In the most widely used method, some of the satellite nodes in the forwarding path can independently calculate the best node for the next hop and send data, and other satellite nodes use the centralized routing algorithm to distribute the routing table to other nodes after calculation.

The disadvantages of the hybrid routing algorithm are its higher complexity, the need for a stable inter-satellite laser link, and higher single-satellite on-satellite processing power. Hybrid routing has become a popular topic in satellite-routing algorithms, owing to its superior network performance.

Summary

As a new communication type, satellite laser communication is characterized by a strong anti-jamming ability, high transmission rate, and large communication capacity. Worldwide research on satellite laser communication networks has mainly focused on satellite networking and inter-satellite transmission. As research on satellite laser communication networks continues to deepen, the technology will become more mature, and its application prospects will become increasingly broad.

This chapter provided an overview of the current status of domestic and international research on satellite laser communication networks, described the key technologies of satellite laser communication, and briefly introduced satellite laser communication networks.

Satellite laser communication networks have broad development prospects, and further developments in the future can be carried out, based on the following points:

1. Reduce the impact of dynamic network-topology changes on network performance. Topological changes in satellite laser networks seriously affect the network performance; for example, they cause frequent inter-satellite laser link interruptions and increase the routing-update frequency. Therefore, satellite laser networks must introduce flexible and dynamic solutions to minimize the impact of dynamic changes in the satellite laser network topology.
2. Rational use of satellite laser communication network resources. Link and hardware resources are scarce in satellite laser communication networks, and maintenance is also a problem. Therefore, improving the utilization of network-link resources and reducing the consumption of on-satellite hardware resources are urgent problems that must be solved.
3. Multi-service bearing. The purpose of a satellite laser network is to provide wide coverage and seamless, all-round, diverse, high-quality, and efficient communication services to users worldwide. Therefore, suitable technology is required to carry multiple transmission services in the network.

References

1. Jiang Huilin, Fu Qiang, Zhao Yiwu et al. Status and trends in the development of space information networks and laser communications [J]. Journal of Internet of Things, 2019, 3(02):1–9.
2. Vincent W. S. Chan. Optical Satellite Networks[J]. Journal of Lightwave Technology, 2003, 21(11):2811–2828.
3. Nikos Karafolas, Stefano Baroni. Optical Satellite Networks[J]. Journal of Lightwave Technology, 2000, 18(12):1792.
4. Tolga Ors, Catherine Rosenberg. Providing IP QoS over GEO satellite systems using MPLS [J]. International Journal of Satellite Communications, 2001, 19(5):443–461.
5. Ekici E, Akyildiz I F, Bender M D. Datagram Service Distributed Routing Algorithm Strategy in LEO Satellite Communication Network[J]. IEEE/ACM Trans. on Networking, 2001, 9(2): 137–148.

6. Anton Donner, Matteo Berioli, Markus Werner. MPLS-based satellite constellation networks [J]. IEEE Journal on Selected Areas in Communications, 2004, 22(3):438–449.
7. Papapetrou E, Karapantazis S, Pavlidou F N, et al. Multiservice on-demand routing in LEO satellite networks[J]. IEEE Transactions On Wireless Communications, 2007, 8(1): 107–113.
8. Papapetrou E, Karapantazis S, Pavlidou F N. Distributed on-demand routing for LEO satellite systems[J]. Computer Networks, 2007, 51(15):4356–4376.
9. Oemer Korcak, Fatih Alagoez, Abbas Jamalipour. Priority-based adaptive routing in NGEO satellite networks[J]. International journal of communication systems, 2007, 20(3):313–333
10. Fu X, Sun L J, Ye X G, et al. Routing algorithm for MPLS traffic engineering in satellite network [J]. Journal on Communications, 2011, 32(5):104.
11. Kawamoto, Y, Nishiyama, et al. A Traffic Distribution Technique to Minimize Packet Delivery Delay in Multilayered Satellite Networks[J]. IEEE Transactions on Vehicular Technology, 2013, 62(7):3315–3324.
12. Pan C, Du H, Liu Q. A routing algorithm for mpls traffic engineering in leo satellite constellation network[J]. Int. J. Innov. Comput. Inf. Control, 2013, 9(10):4139–41410.
13. Sushank C, Angela A, Kashif N.Realization of free space optics with OFDM under atmospheric turbulence[J]. Optik, 2014, 125(18):5196–5199.
14. R. Ferrús, H. Koumaras, O. Sallent, G. Agapiou, T. Rasheed, M.-A. Kourtis, C. Boustie, P. Gélard, T. Ahmed. SDN/NFV-enabled satellite communications networks: Opportunities, scenarios and challenges[J]. Physical Communication, 2016, 18(2):95–113.
15. Huang J, Su Y, Huang L, et al. An Optimized Snapshot Division Strategy for Satellite Network in GNSS [J]. IEEE Communications Letters, 2016, 20(12): 2406–24010.
16. Georgios Gardikis, Harilaos Koumaras, Chris Sakkas, Vaios Koumaras. Towards SDN/NFV-enabled satellite networks[J]. Telecommunication Systems, 2017, 66(4):615–62.
17. Lyras Nikolaos K., Kourogiorgas Charilaos I., Panagopoulos Athanasios D. Cloud Free Line of Sight Prediction Modeling for Optical Satellite Communication Networks[J]. IEEE Communications Letters, 2017, 21(7):1537–1540.
18. Khallaf H S, Shalaby H M H, Garrido-Balsells J M, et al. Performance analysis of a hybrid QAM-M PPM technique over turbulence-free and gamma–gammafree-space optical channels [J]. Journal of OpticalCommunications and Networking, 2017, 9(2):161–171.
19. Lyras Nikolaos K., Efrem Christos N., Kourogiorgas Charilaos I., Panagopoulos Athanasios D. Optimum Monthly Based Selection of Ground Stations for Optical Satellite Networks [J]. IEEE Communications Letters, 2018, 22(6):1192–1195.
20. Networks - Telecommunications; Joint Placement of Controllers and Gateways In Sdn-enabled 5g-satellite Integrated Network[J]. Telecommunications Weekly, 2019, 36(2):221–232.
21. Wang F, Jiang D, Qi S. An adaptive routing algorithm for integrated information networks [J]. China Communications, 2019, 12(02): 108–132.
22. Kongzhe Yang, Bangning Zhang, Daoxing Guo. Partition-Based Joint Placement of Gateway and Controller in SDN-Enabled Integrated Satellite-Terrestrial Networks[J]. Sensors, 2019, 19(12):2774.
23. Lyras Nikolaos K., Efrem Christos N., Kourogiorgas Charilaos I., Panagopoulos Athanasios D., Arapoglou Pantelis. Daniel. Optimizing the Ground Network of Optical MEO Satellite Communication Systems[J]. IEEE Systems Journal, 2019, 14(3):3968–3976.
24. Liu W, Tao Y, Liu L. Load-Balancing Routing Algorithm Based on Segment Routing for Traffic Return in LEO Satellite Networks [J]. IEEE Access, 2019, 7(01): 112044–112053.
25. Liu J, Luo R, Huang T, et al. A Load Balancing Routing Strategy for LEO Satellite Network [J]. IEEE Access, 2020, 8(01): 155136–155144.
26. Polnik M, Arulselvan A, Riccardi A. Scheduling space-to-ground optical communication under cloud cover uncertainty [J].IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(5):2838–28410.
27. S. Chan and V. W. S. Chan. Constellation topologies for a space-based information network backbone using optical intersatellite links[J]. IEEE Access, 2004, 2(01):812–821:

28. Li Yongjun, Zhao Shanghong, Zhang Dongmei, et al. 空Laser communication link networking technology for inter-fleet satellite platforms [J]. Optical Communication Technology, 2006(10): 47–410.
29. Yi Xianqing, Feng Mingyue, Zhao Yang, et al. A Study on Destruction Resistant Routing for Satellite Networks Based on GEO/MEO Star Layer Networking [J]. computer science, 2007(08): 74–82.
30. Zou Qi, Bai Yuebin, Wu Wei, et al. An efficient routing algorithm for LEO satellite constellations [J]. High-tech communications, 2008, 18(08):795–800.
31. Qinglong Yang, Liying Tan, Jing Ma. Analysis of Crosstalk in Optical Satellite Networks With Wavelength Division Multiplexing Architectures[J]. Journal of Lightwave Technology, 2010, 28(6):931–939.
32. Li Yongjun, Zhao Jihong, Wu Jili, et al. Global Coverage Stable Topology LEO/MEO Dual Layer Satellite Laser Network Design [J]. Optoelectronics. Laser, 2009, 20(03):324–329.
33. Tan L, Yang Q, Ma J, et al. Wavelength Dimensioning of Optical Transport Networks Over Nongeosychronous Satellite Constellations[J]. Journal of Optical Communications and Networking, 2010, 2(4):166–174.
34. Yang, Qinglong, Tan, Liying, Ma, Jing, Jiang, Shiqi. An Analytic Method of Dimensioning Required Wavelengths for Optical WDM Satellite Networks[J]. IEEE communications letters: A publication of the IEEE Communications Society, 2011, 15(2):247–2410.
35. Li Yong Jun, Wu Ji Li, Zhao Shang Hong, Meng Wen, Ma Li Hua, Shi Lei, Chu Xing Chun, Hou Rui, Li Tian. A novel two-layered optical satellite network of LEO/MEO with zero phase factor[J]. Science China Information Sciences, 2010, 53(6):1261–1276.
36. Gao Lijuan, Jiang Taijie. Analysis and Improvement of Satellite Network Connectivity and Efficient Routing Algorithms [J]. Systems Engineering and Electronics, 2014, 36(10): 2071–2075.
37. Yang Li, Li Jingsen, Wei Debin. A congestion control algorithm for highly dynamic satellite networks [J]. Astronautical Journal, 2014, 35(08):953–960.
38. Yi Dong, Shanghong Zhao, Hao dan Ran, Yongjun Li, Zihang Zhu. Routing and Wavelength Assignment in a Satellite Optical Network Based on Ant Colony Optimization With the Small Window Strategy[J]. Journal of Optical Communications and Networking, 2015, 7(10): 995–1000.
39. Chen Erhu, Liu Ying, Liu Lu, et al. Network design and performance analysis of laser link relay satellite systems [J]. Infrared and Laser Engineering, 2015, 44(09):2778–2782.
40. Chen Jianfei. A Study of Routing Strategies for Multi-Layer Satellite Networks [D]. Harbin. Harbin Institute of Technology, 2018.
41. Zheng Yongxing, Zhao Shanghong, Liu Yun, Li Yongjun, Tan Qinggui, Xin Ning. Weighted Algebraic Connectivity Maximization for Optical Satellite Networks[J]. IEEE Access, 2017, 5(99):6885–6893.
42. Yongxing Zheng, Shanghong Zhao, Yun Liu, Qinggui Tan, Yongjun Li, Yong Jiang. Topology control in self-organized optical satellite networks based on minimum weight spanning tree [J]. Aerospace Science and Technology, 2017, 69(10):449–458.
43. Guoli Wen, Qi Zhang, Houtian Wang, Qinghua Tian, Ying Tao. An Ant Colony Algorithm Based on Cross-Layer Design for Routing and Wavelength Assignment in Optical Satellite Networks[J]. China Communications, 2017, 14(08):63–75.
44. Xu Shuang, Wang, Xingwei, Huang Min, et al. A satellite network capacity intelligent analysis method [J]. Journal of Computing, 2017, 40(07):1572–1582.
45. Li Tingting. Research on resource scheduling mechanism of space-based backbone network based on optical burst switching [D]. Beijing. Beijing University of Posts and Telecommunications, 2019.
46. Liu Yeqi, Zhang Qi, Xin Xiangjun, Tian Qinghua, Tao Ying, Liu Naijin, Lv Kai. Bee Colony Algorithm Optimization Based on Link Cost for Routing and Wavelength Assignment in Satellite Optical Networks[J]. IEICE Transactions on Communications, 2019, 28(6):690–702.

47. Xianfeng Liu, Xiaoqian Chen, Lei Yang, Quan Chen, Jianming Guo, Shuai Wu. Dynamic topology control in optical satellite networks based on algebraic connectivity[J]. Acta Astronautica, 2019, 165(12):287–298.
48. Zhang Jiale, Qian Hongyan, Cheng Xiang, et al. A multi-beam switching mechanism for software-defined satellite networks [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(12):2503–25010.
49. Xianfeng Liu, Lei Yang, Quan Chen, Jianming Guo, Shuai Wu, Xiaoqian Chen. An analytic method of wavelength requirements in dynamic optical satellite networks[J]. IEEE Communications Letters, 2020, 24(11):2569–2573.
50. Cong Peng, Shanghong Zhao, Ruixin Li, Jun Li, Weilong Wang, Hanghang Gao. Separated traffic grooming for Low Earth Orbit (LEO) optical satellite networks with wavelength dimensioning[J]. International Journal of Satellite Communications and Networking, 2020, 38(6): 499–512.
51. Xue Sun, Suzhi Cao. A Routing and Wavelength Assignment Algorithm Based on Two Types of LEO Constellations in Optical Satellite Networks[J]. Journal of Lightwave Technology, 2020, 38(8):2106–2114.
52. Zhang Lu. Research on Routing Algorithms for Low and Medium Orbit Satellite Communication Networks [D]. Nanjing: Southeast University,2021.
53. Wang Ruisong, Zhong Zhicong, Liu Gongliang, et al. Network coding based capacity enhancement method for satellite networks [J]. mobile communications, 2021, 45(05):2–8.
54. Liu Yue, Huang Yin, He Jiangtao, et al. A step-tracking method for multibeam satellite communication networks[J]. Telecommunication Technology, 2022, 62(12):1734–1740.
55. Zuo Peiliang, Wang Chen, Jiang Hua. A Reinforcement Learning-based Intelligent Routing Algorithm for Low-Orbit Satellite Networks[J]. Journal of Beijing Institute of Electronic Science and Technology, 2022, 30(02):35–43.
56. Zheng Jianyang, Li Hui, Zhou Yuling et al. A traffic load balancing algorithm for star-earth networks[J]. Telecommunication Technology, 2022, 62(08):1085–1092.
57. Liu Yuting, Tian Feng, Li Guotong. An Efficient Reverse Seam Link Building Strategy for Low Orbit Satellite Networks[J]. Journal of Chinese Academy of Sciences University, 2023, 40(02): 240–2410.
58. Xie Renchao, Yang Yutian, Tang Qinqin et al. On-board edge DNN inference strategy for low-orbit satellite networks[J]. Journal of Beijing University of Posts and Telecommunications, 2023, 46(02):57–63.
59. Gagliardi RM, Karp S. Optical communications [M]. New York, NY, USA: Wiley, 1976.
60. Qiang Fu, Huilin Jiang, Xiaoman Wang, et al. Current status and development trend of space laser communication research [J]. China Optical, 2012, 5(2): 116–125.
61. Laurent Franck, Gérard Maral. Static and adaptive routing in ISL networks from a constellation perspective[J]. International Journal of Satellite Communications, 2002, 20(6):455–475.
62. Chen Weiqi. Research on Satellite Link Optimization and Analysis Techniques for Multi-Layer Satellite Networks [D]. Beijing: Beijing University of Posts and Telecommunications, 2019
63. Han Huipeng. An analysis of interstellar links in low orbit communication constellations [J]. Satellite and Networks. 2018, 0(8):40–42
64. Wang Zhenyong. Design and Analysis of Multi-Layer Satellite Network Structure [D]. Harbin: Harbin Institute of Technology, 2008.
65. Jiang Chang, Fan Xiaolin. A preliminary study of aerospace laser communication and tracking (C&T) technology (below) [J]. Telemetry, 2005, 26(5):18–23.
66. Franck L, Maral G. Static and adaptive routing in ISL networks from a constellation perspective [J].International Journal of Satellite Communications, 2002, 20(6): 455–475.
67. NI Shaojie, YUE Yang, ZUO Yong, et al. Current status and outlook of satellite network routing technology[J]. Journal of Electronics and Information, 2023, 45(02):383–395.

Pulse-Like Position Modulation Technology

3

Contents

Introduction 46
Foreign Research Progress 47
Domestic Research Progress 49
Pulse-Position Modulation Technology 51
Symbol-Structure Analysis of Various Modulation Modes 52
Performance Analysis of the Modulation Modes 52
Slot Error Rate Analysis 58
Channel Capacity 63
Adaptive Threshold Detection 74
Factors Affecting the Laser-Modulation Rate 78
Summary 80
References 81

Abstract

Modulation and demodulation are key technologies to improve the transmission efficiency of optical wireless communication. Different modulation methods have different performances. This chapter summarizes the research progress of various types of pulse-position modulation at home and abroad and then introduces the research from the Xi'an University of Technology in the area of pulse-position modulation. For on-off keying, pulse-position modulation, multiple pulse-position modulation, differential pulse-position modulation, overlapping pulse-position modulation, dual duration position modulation, dual-amplitude pulse-position modulation, digital pulse interval modulation, double-headed pulse interval modulation, dual-pulse interval modulation, dual-amplitude pulse interval modulation, fixed-length digital pulse interval modulation, fixed-length dual-amplitude pulse interval modulation, shorten pulse position modulation and separated double-pulse position modulation, the performance of the symbol structure, bandwidth requirement, average transmit power, time-slot error rate, and average channel capacity are analyzed and compared. Finally, the development direction of pulse-like position modulation is prospected.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_3

Keywords

Optical wireless communication · Modulation and demodulation · Average channel capacity · Bandwidth requirement · Bit-error rate

Introduction

Modulation and demodulation are key technologies for improving the transmission efficiency of optical wireless communication. Different modulation methods exhibit different performance. Pulse-like position modulation refers to pulse-position modulation (PPM), differential pulse-position modulation (DPPM), and their combination, as well as various pulse-position modulation modes that have evolved from them. The research progress of pulse-like position modulation at home and abroad is reviewed for the following modes:

1. On-off keying
2. Pulse-position modulation
3. Multiple pulse-position modulation
4. Differential pulse-position modulation
5. Overlapping pulse-position modulation
6. Dual-duration pulse-position modulation
7. Dual-amplitude pulse-position modulation
8. Digital pulse-interval modulation
9. Double-headed pulse-interval modulation
10. Dual-pulse pulse-interval modulation
11. Dual-amplitude pulse-interval modulation
12. Fixed-length digital pulse-interval modulation
13. Fixed-length dual-amplitude pulse-interval modulation
14. Shortened pulse-position modulation
15. Separated double-pulse position modulation

In addition, their symbol structure, bandwidth requirements, average transmit power, time-slot error rate, and average channel capacity are analyzed and compared.

When a laser beam is transmitted through an atmospheric channel, the error-control performance of the optical communication system deteriorates, owing to various factors, such as atmospheric attenuation and atmospheric turbulence. Researchers have studied a series of modulation techniques to improve communication efficiency. An advanced modulation format and corresponding demodulation technology should be selected to ensure efficient transmissions and a low bit-error rate (BER) to reduce the interference of the atmospheric environment on the communication system. Therefore, research on modulation modes is an important topic in the field of optical wireless communication.

Foreign Research Progress

In laser communications, PPM can achieve the highest data-transmission rate with the minimum average optical power for a given laser-pulse repetition rate [1]. Of course, it has some disadvantages, such as increasing the demand for bandwidth, although it improves the anti-interference ability. After decades of development, foreign researchers have studied many types of pulse-position modulation based on PPM. Different modulation methods exhibit different performance in different applications [2].

In 1978, Pierce of the US Jet Propulsion Laboratory (JPL) proposed PPM and applied it to optical wireless communication [3]. In 1982, Prati et al. extended PPM coding to block-pulse coding on multiple PPM frames and proposed a maximum-likelihood demodulation method for the optical-receiver model of shot and thermal noise [4]. In 1983, Prati et al. studied PPM-modulated laser-pulse expansion and discussed several methods for suppressing the expansion, including pulse equalization, extended pulse integration, and pulse-shape matching [5].

In 1983, Garrett et al. used the matched-filter method to analyze the reception sensitivity of a PPM transmission system. The results showed that the reception sensitivity could also be improved using direct detection [6]. In 1986, Gardner et al. analyzed the BER performance of a phase-locked synchronous PPM system and compared it to that of a fully synchronized PPM system. The results showed that when the loop bandwidth is less than 0.1% of the slot frequency, the signal power required to compensate for incomplete synchronization increases only slightly [7].

In 1987, Georgiades et al. derived the maximum-likelihood estimation of PPM symbol synchronization without time-slot synchronization through observing the noise of random modulation data through an M-symbol-wide window [8]. In 1988, Davidson et al. built a 25-Mbit/s direct detection optical communication system using a PPM signal and measured its performance under laboratory conditions [9, 10]. In 1988, Zwillinger et al. used an alternative differential pulse-position modulation (DPPM) coding scheme. The results showed that, in some cases, DPPM has a higher communication rate than PPM in an optical channel with limited noise-free bandwidth and limited average power [11].

In 1989, Sugiyama et al. reported that multi-pulse-position modulation could reduce the transmission bandwidth required by PPM to approximately half, under the same transmission efficiency, thus, improving the bandwidth utilization [12]. In 1994, Barry et al. proposed a method to detect intersymbol interference (ISI) and additive Gaussian white noise when using digital pulse-position modulation. The performance of the recursive equalization scheme was analyzed, and the results showed that the equalizer was effective in mitigating ISI [13].

In 1996, Malik et al. evaluated the performance of PPM on an ISI channel. The BER performance of the unbalanced system was analyzed, and the performance of PPM maximum-likelihood sequence detection (MLSD) in ISI with additive Gaussian white noise was discussed. When MLSD was used, 16-PPM exhibited the best average power efficiency [14].

In 1997, David et al. analyzed the performance of trellis codes using PPM on indoor wireless channels and proved the effectiveness of MLSD in suppressing multipath phenomena. They also evaluated the performance of 8-PPM with a 2/3 encoding rate and 16-PPM with a 3/4 encoding rate on multipath channels [15].

In 1998, Park et al. proposed a trellis code that used multi-pulse-position modulation (MPPM). Compared with trellis codes using PPM, the code was less sensitive to multipath phenomena and provided better power efficiency when the expected bit rate was greater than the channel bandwidth [16]. In 1999, Kahn et al. studied the expressions of the error probability and power spectral density of DPPM in the presence of Gaussian white noise [17].

In 2001, Nikolai et al. studied the use of chaotic pulse-position modulation (CPPM) to modulate binary information to the carrier through the pulse position. They discussed the basic design principles of a chaotic pulse generator, synchronization, and the performance of a chaotic pulse communication scheme in the presence of channel noise [18].

In 2002, Aldibbiat et al. studied the performance of a double-headed pulse-interval modulation (DHPIM) scheme in an indoor multipath light environment. The results showed that DHPIM and PPM required slightly higher optical power than digital pulse-interval modulation (DPIM) [19]. In 2005, Moision et al. proposed serial cascaded pulse-position modulation (SCPPM), which is composed of the serial concatenation of an external convolutional code, interleaver, bit accumulator, and PPM and is encoded by a shift-register operation and table lookup [20].

In 2005, Kiasaleh et al. analyzed the error rate expression and the upper limit of the PPM free-space optical (FSO) communication system when an avalanche photodiode (APD) detector was used to detect the light field [21]. In 2005, Wilson et al. studied a method that combined multiple laser transmitters with multiple photodetectors. A Q-ary PPM mode was adopted. For Rayleigh-fading and lognormal-fading channel models, multiple-input/multiple-output (MIMO) transmission was adopted to suppress the fading effect on turbulent channels [22].

In 2006, Ivan et al. proposed Q-ary PPM for MIMO transmission. It was composed of multistage coding and a low-density parity-check code. This scheme yielded good coding gains under strong turbulence conditions [23].

In 2007, Maged et al. studied and compared two PPM-based iterative-coded modulation techniques for deep-space optical communications. The first encoding was represented by serially concatenated SCPPM, which involved the serial concatenation of an external convolutional code, interleaver, bit accumulator, and PPM. The second was a serially concatenated code, in which the constituent code was composed of a low-density even parity-check (LDPC) code and serially concatenated LDPC-PPM code [24].

In 2008, Sethakaset et al. proposed a concatenated decoding algorithm for hard and soft decision detection. The results showed that when the channel was non-dispersive, the coding gain provided by the concatenated code was about 4 dB. On dispersive channels, when the ratio of the delay spread to bit duration was low, the performance of the coding system was better than that of non-coding systems [25].

In 2009, Bounnane et al. analyzed the use of classical binary convolutional codes, iterative soft deconvolution, and channel decoding in receivers. The influence of bit-symbol mapping on the performance of iterative receivers was studied, and an MPPM mapping design rule was proposed [26].

In 2012, Wasiu et al. studied a modulation scheme that combined spatial shift keying (SSK) and PPM. Shortened pulse-position modulation (SPPM) is a multi-transmitter scheme that combines the simplicity and spectral efficiency of SSK with the energy efficiency of PPM technology [27]. In 2016, Correa et al. carried out numerical simulations of the encoding, decoding, and propagation performance of short pulses and words in on-off keying (OOK), pulse-amplitude modulation (PAM), and PPM modulation modes in optical code-division multiple-access (OCDMA) systems [28].

In 2017, Zohaib et al. discussed the implementation of a binary pulse-position modulation (DuoPPM) coding scheme and its application to visible-light communication systems. The experimental results achieved at a 14 Mbit/s data rate showed that the bit-error rate was less than 10^{-10}. The practicability, robustness, and limitations of DuoPPM were analyzed [29].

In 2020, Sharma et al. studied a free-space optical communication system using a hybrid modulation technology including pulse-position modulation, binary phase-shift keying, and subcarrier-intensity modulation. Lognormal and double-gamma probability-density functions were used as channel-fading statistics under weak, medium, and strong turbulence conditions. The average bit-error rate and closing expression of the hybrid modulation scheme were derived [30].

In 2020, Zaiton et al. studied the performance of two single-carrier modulation schemes for visible-light communication (VLC), namely, OOK and PPM. The results showed that PPM is superior to OOK in terms of transmission distance and path loss, and OOK is the best choice for achieving the highest data rate [31].

Since 1978, research on pulse-like position modulation has focused on improving the communication quality from the initial performance analysis of single PPM and DPPM modulation [4–7] to methods cascading PPM with other coding methods [24–27]. The research content has also developed from theoretical derivations and simulations [3–5] to practical applications [9, 10]. Communication systems have been developed from single-input single-output (SISO) communication systems [18–20] to multiple-input multiple-output (MIMO) communication systems [22, 23]. MIMO can reduce the fading effect on turbulent channels. The bit-error rate performance [7, 21], reception sensitivity [6], frequency-band utilization [12, 17], intersymbol interference [13, 14], and other aspects are analyzed.

Domestic Research Progress

In 1998, Wang et al. studied a time-slot synchronizer using a digital phase-locked loop to realize PPM. The anti-noise performance of the time-slot synchronizer using this method for loop filtering was better than that using the mean, three-point, and screening methods. With other performance indicators unchanged, this algorithm

ensured that the entire synchronization system worked reliably under stronger background noise [32].

In 1999, Huang et al. designed a PPM modulator. Digital circuits, such as counting, latch, and control, were realized using programmable logic devices, which greatly simplified the hardware circuit [33].

In 2000, Zhang et al. proposed a new solution using a digital-logic design for slot synchronization, word synchronization, and word synchronization in the frame synchronization of pulse-position modulation [34].

In 2000, Zou et al. compared the PPM, OOK, and PAM modulation modes in terms of the average optical power and bandwidth utilization. This showed that PPM had a high energy efficiency, at the cost of reducing the bandwidth [35].

In 2000, Zhang et al. analyzed PPM, DPPM, and MPPM, compared them with OOK, and analyzed their advantages and disadvantages [36].

In 2002, Pang et al. analyzed and compared the power utilization ratio, frequency-band utilization ratio, transmission capacity, and anti-symbol interference performance of OOK, PPM, DPPM, and DPIM in optical wireless communication [37].

In 2005, Tang et al. studied the error performance of digital pulse-interval modulation (DPIM), analyzed the symbol structure, transmission power, bandwidth requirements, and other metrics, and compared them with OOK and PPM [38].

In 2006, Zhang et al. designed a PPM optical wireless communication system based on a field-programmable gate array (FPGA), based on the study of PPM modulation technology [39].

In 2007, Wang et al. analyzed and compared the bandwidth requirements, transmission capacity, power requirements, and error performance of OOK, PPM, and DPIM, based on their modulation structures, combined with the characteristics of optical wireless channels [40].

In 2010, Sha et al. used LDPC codes as a channel code, combined with a simplified software information-extraction method, to simulate an LDPC-encoded PPM system in an atmospheric turbulence channel. Simulation results showed that the scheme could meet the requirements of optical wireless communication systems [41].

In 2010, Cheng et al. presented dual-pulse pulse-interval modulation (DPPIM). They analyzed the bandwidth requirement, transmission capacity, and average transmission power, deduced the packet error rate model under a weak turbulence channel, and compared it with PPM, DPIM, fixed-length digital pulse-interval modulation (FDPIM), OOK, and other methods [42, 43].

In 2016, Xiang et al. proposed a method to convert the soft information output by SCPPM decoding into the expected value of each time slot of PPM. The SCPPM decoding output first makes a hard decision and then obtains the accumulator state through the accumulation module. Depending on this state, the next PPM mapping mode is selected to obtain the probability and expectation of each PPM time slot, and the best sampling point is obtained by predicting the clock deviation [44, 45].

In 2019, Liu et al. studied an end-to-end RF–FSO hybrid system using a hybrid pulse-position modulation–binary phase-shift keying–subcarrier-intensity modulation (PPM–BPSK–SIM) scheme. Radio-frequency (RF) links follow a Rayleigh

distribution and FSO links follow a double-gamma distribution. Considering the influence of atmospheric turbulence and pointing-error conditions, the average bit-error rates of various PPM–BPSK–SIM schemes were derived, and the outage probability and average channel capacity of the system were discussed [46].

In 2020, Qinling et al. analyzed and compared the performance of visible-light communication systems based on OOK, PPM, and DPIM. The coding structure, bandwidth requirements, and average transmission power of OOK, PPM, and DPIM were compared [47].

The study of pulse-like position modulation has developed from the study of PPM performance [32–34] to the comparative development of multiple modulation performance [35–38]. Various modulation methods are related to one another. Each modulation method is proposed, based on the shortcomings of existing modulation methods. However, the new modulation methods also have shortcomings. To date, nearly 20 types of pulse-like position modulation have been proposed. However, these 20 modulation methods still have unsolved problems. In future research on atmospheric laser-communication systems, additional modulation methods will be proposed to improve the communication-system performance.

Pulse-Position Modulation Technology

Professor Ke Xizheng of Xi'an University of Technology and others have studied optical wireless communication technology since 2000 and have made good progress in signal modulation, coding, antennas, and automatic tracking and aiming technologies [1]. Atmospheric laser-communication systems, including serial ports, parallel ports, Ethernet interfaces, and other common digital interfaces, have been successfully developed [2]. A PPM modulation and demodulation system was designed. Time-slot synchronization and frame synchronization during demodulation were realized using phase-locked loop technology and Gold codes, respectively. A time-slot synchronization technology was invented [72].

Professor Ke's team comprehensively analyzed and compared the performance of more than ten modulation modes and determined the best modulation modes for certain applications. Matching modulation modes were selected for different channels and the shortcomings of these modulation modes were determined to provide a theoretical basis for selecting the appropriate modulation mode in different situations.

To choose a suitable modulation mode, the research group of Xi'an University of Technology carried out a systematic analysis of pulse-like position modulation modes. In 2005, Ding et al. [48] analyzed the performance of PPM and implemented it using hardware. In 2006, Zhao [55] studied DPPM and analyzed its performance. In 2007, Qinling et al. [49] studied the symbol structure and constellation of MPPM and implemented it using hardware. In 2007, Huang [56] analyzed the basic principles and performance of overlapping pulse-position modulation (OPPM).

In 2010, Ma [54] proposed a new modulation technology for atmospheric laser communication that combined PPM and polarization modulation [73]. In 2011,

Sun [57] analyzed the symbol structure, power spectral density, bandwidth requirements, etc. of dual-amplitude pulse-position modulation (DAPPM). In 2011, Jia [58] analyzed the performance of DHPIM. In 2013, Liu [59] systematically studied the symbol structure, power spectral density, bit-error rate, etc. of pulse-like position modulation.

In 2016, Shi [60] studied quadrature amplitude modulation (QAM) and demodulation programs and analyzed the results. In 2015, Tian [61] discussed an adaptive threshold-detection model and analyzed its performance. In 2012, Pei [62] studied two methods of suppressing turbulence and improving communication performance: multibeam emission and multiaperture reception. In 2019, Liu [63] studied and optimized the response characteristics of semiconductor lasers and photodetectors to improve the modulation rate of communication systems.

Symbol-Structure Analysis of Various Modulation Modes

The main modulation modes of intensity modulation/direct detection systems include OOK, PPM, DPIM, and DHPIM. These modulation modes use a laser as the carrier to transmit information. The receiving end of the transmission determines the transmitted information, based on the position of the optical pulse in an equally spaced time period; this can be collectively referred to as pulse-like position modulation.

Taking the modulation bit number $M = 4$ as an example, the symbol structure of each modulation mode is shown in Fig. 3.1 [53]. In Fig. 3.1, α represents the pulse-width parameter and β represents the amplitude parameter. Among the above code words, fixed-length digital pulse-interval modulation (FDPIM) has the longest symbol length, followed by fixed-length dual-amplitude pulse-interval modulation (FDAPIM), PPM, and dual-amplitude pulse-interval modulation (DAPIM). The shortest is OOK.

Performance Analysis of the Modulation Modes

In optical wireless communication systems, OOK modulation was the earliest used in the intensity modulation/direct detection systems; however, its performance is lacking in various aspects, such as anti-interference ability and power utilization. To improve these metrics, modulation modes, such as PPM [48], DPIM [58], and DHPIM [58], have been proposed. Next, we discuss some basic performance metrics of these modulation methods, such as the average symbol length, bandwidth requirements, and average transmission power.

Average Symbol Length

Figure 3.2 shows the average symbol length of the various modulation modes. To clearly distinguish the average symbol-length curve of different modulation modes, the pulse-width parameter α was set to 10 [1]. It can be seen from the figure that

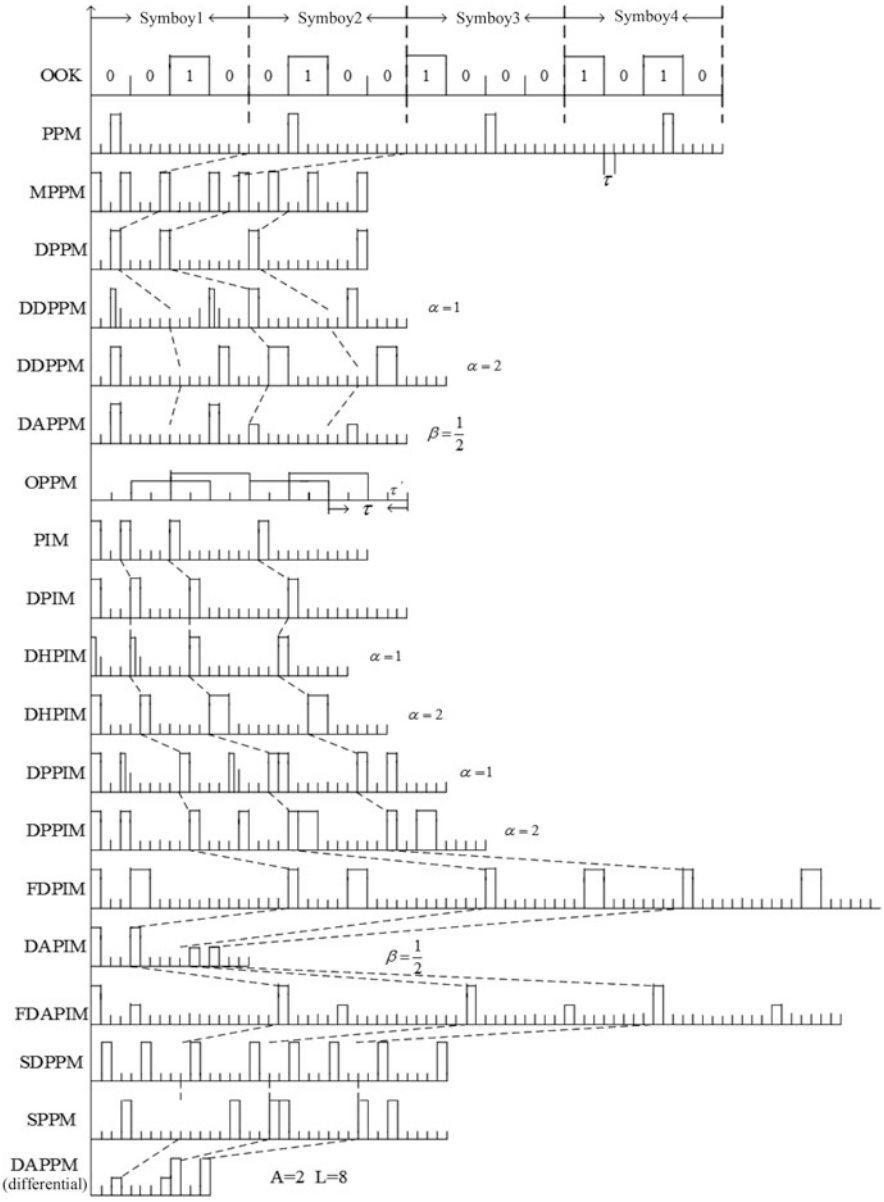


Fig. 3.1 Symbol structure of each modulation mode [53]

when the modulation order M is very small, the average symbol length of these modulation methods makes little difference. With an increase in M, the average symbol lengths of FDPIM, fixed-length dual-amplitude pulse-interval modulation

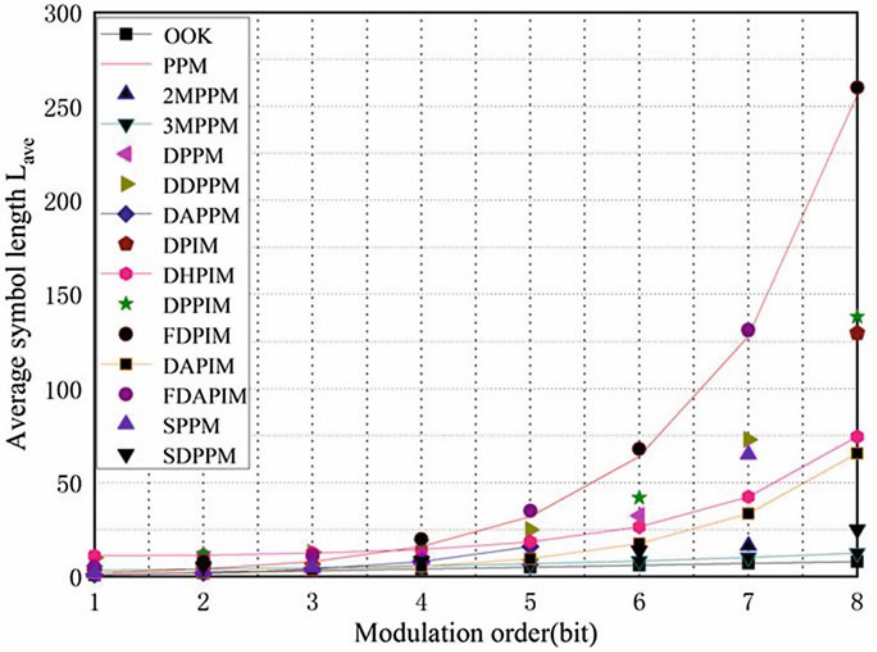


Fig. 3.2 Average symbol length of each modulation mode [53]

(FDAPIM), and PPM increase the fastest. They are longer than the other modulation methods, and their symbol lengths are very close.

For modulation modes with parameters, the average symbol length increases with an increase in the parameters. A comparison of the average symbol lengths of different modulation modes through numerical simulations using MATLAB is shown in Fig. 3.2.

Bandwidth Demand and Utilization

A optical wireless communication system needs a certain bandwidth to transmit information, and the smaller the bandwidth, the better. Figures 3.3 and 3.4 show comparisons of the bandwidth requirements and bandwidth utilization of various modulation modes. The pulse-width parameter α was set to 2, and the information bit rate was 1.

As shown in Fig. 3.3, the bandwidth requirement of each modulation mode increases with an increase in the number of modulation bits, and the bandwidth requirement of the modulation mode with a variable pulse width decreases with an increase in parameter α. As shown in Fig. 3.4, the bandwidth utilization of each modulation mode decreases with an increase in the number of modulation bits [2].

Differential-amplitude pulse-position modulation (DAPPM) and OPPM differ from those mentioned above. The relationship between the DAPPM bandwidth requirements and symbol length is shown in Fig. 3.5. The bandwidth requirement

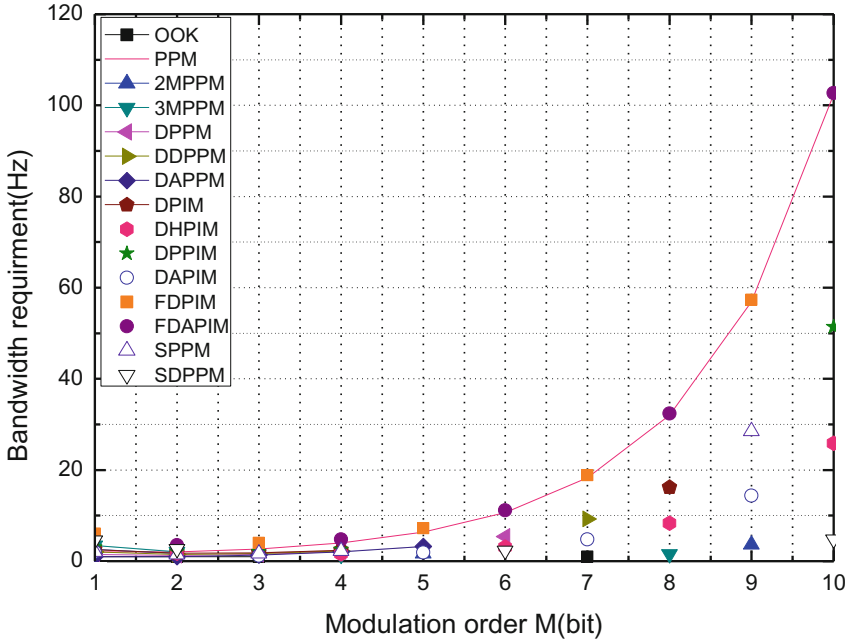


Fig. 3.3 Bandwidth requirements [53]

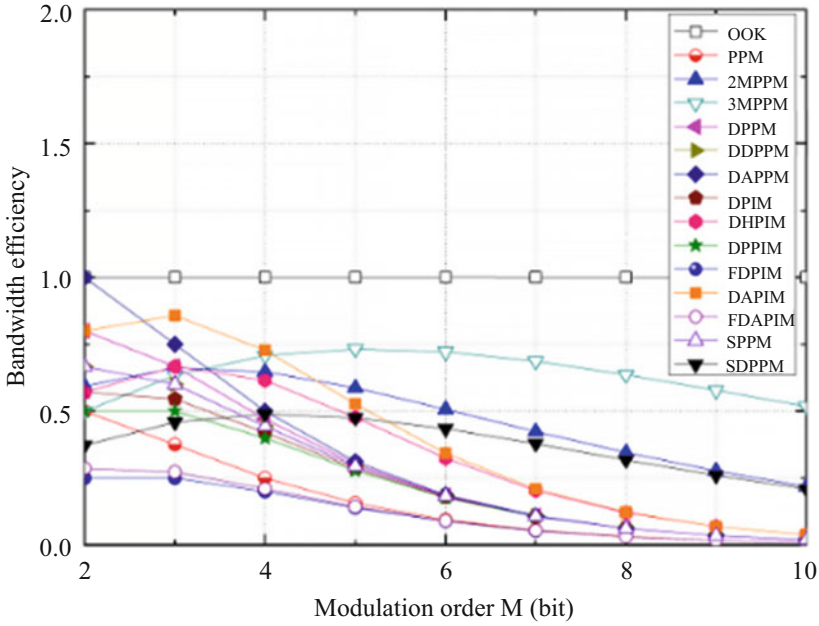


Fig. 3.4 Bandwidth utilization [53]

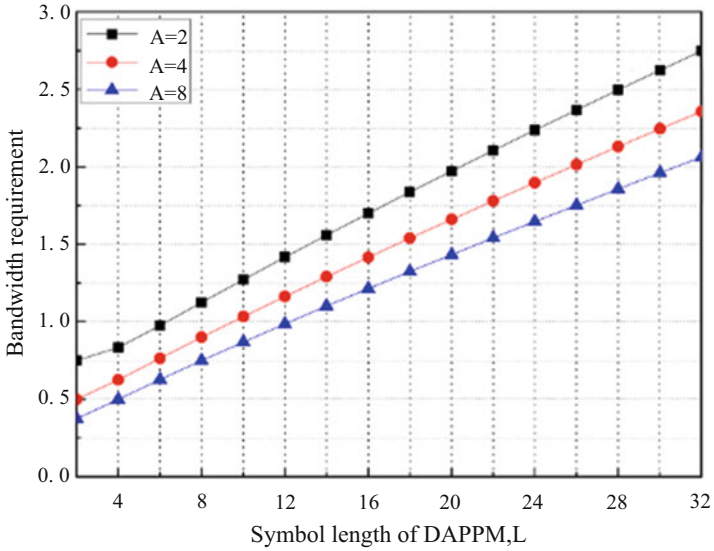


Fig. 3.5 Bandwidth requirements of DAPPM [57]

of DAPPM increases with an increase in the symbol length and decreases with an increase in the pulse amplitude. The relationship between the bandwidth requirement of OPPM and the overlap coefficient is shown in Fig. 3.6. The bandwidth requirement of OPPM decreases with an increase in the overlap coefficient and increases with an increase in the modulation order [50].

Average Transmit Power

In an atmospheric laser communication system, pulse-like position modulation methods can be considered to send "0" and "1" sequences with equal probability. No power is required when sending "0" sequences, and the peak power P_c is required when sending "1" sequences. Therefore, for a given peak-emission power P_c, the average emission power can be simply determined as the probability of sending "1" times the peak power; that is, $P_{\text{ave}} = P_I * P_c$. Figure 3.7 shows the relationship between the average transmission power and modulation order when the peak transmission power is fixed. The pulse-width parameter α was taken as 1, and the amplitude parameter β was taken as 2 [1].

Performance Analysis

The average symbol length, bandwidth requirements, and average transmission power of the various modulation modes are listed in Table 3.1. From the symbol structure and average symbol length of these modulation modes, it is evident that the most important modulation modes are PPM and DPIM. The former represents the transmitted information using the position of the optical pulses in the time slot and

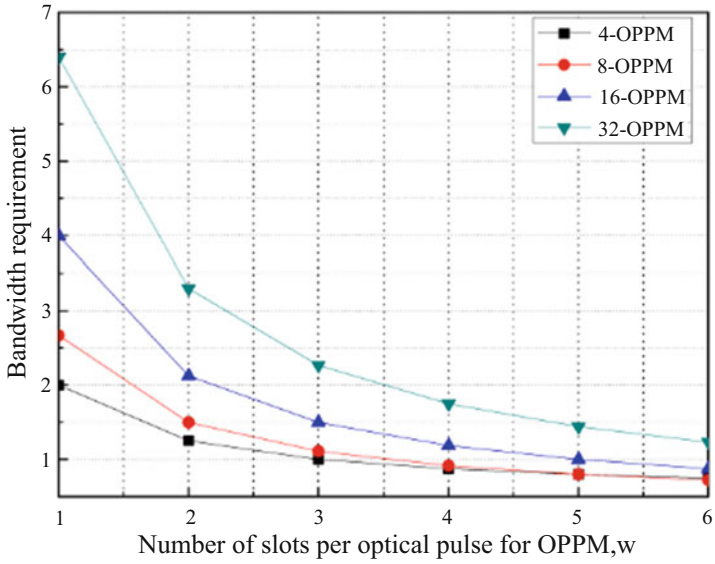


Fig. 3.6 Bandwidth requirements of OPPM [52]

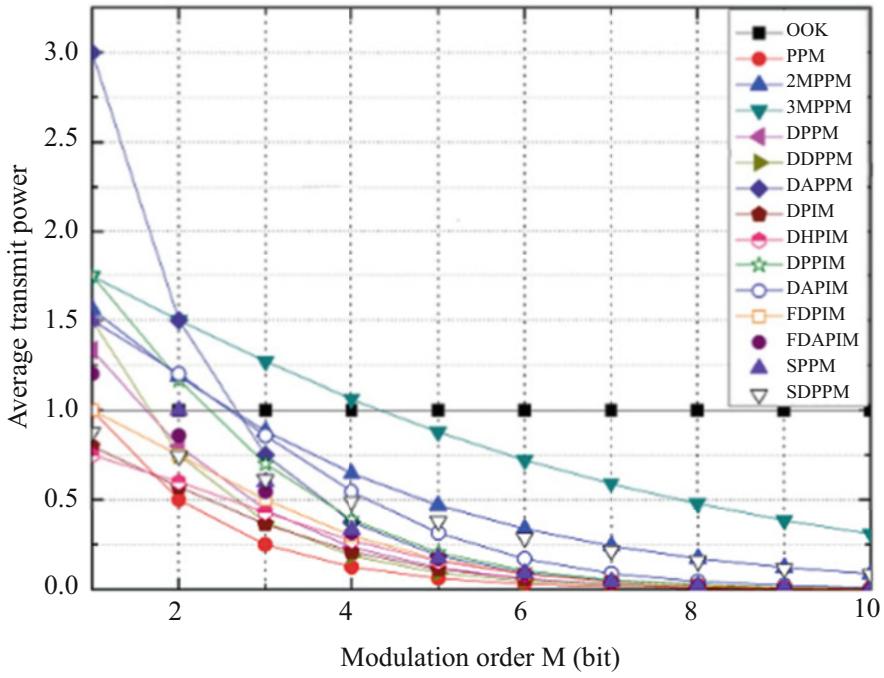


Fig. 3.7 Average transmission power of different modulation methods [53]

Table 3.1 Average symbol length, bandwidth requirements, and average transmission power of each modulation mode

Modulation mode	Average symbol length	Bandwidth requirements	Average transmit power
OOK	M	R_b	$\frac{P_c}{2}$
PPM	2^M	$\frac{2^M R_b}{M}$	$\frac{P_c}{2^M}$
MPPM	n (with p pulses)	$\frac{nR_b}{M}$	$\frac{pP_c}{n}$
DPPM	$\frac{2^M+1}{2}$	$\frac{2^M+1}{2M}R_b$	$\frac{2}{2^M+1}P_c$
DDPPM	$2^{M-1}+\alpha-1$	$\frac{2^M+2\alpha-2}{\alpha M}R_b$	$\frac{3\alpha}{\left(2^{M+1}+4\alpha-4\right)}P_c$
DAPPM	2^{M-1}	$\frac{2^{M-1}}{M}R_b$	$\frac{1+\beta}{2^M}P_c$
PIM	$\frac{2^M+1}{2}$	$\frac{2^M+1}{2M}R_b$	$\frac{2}{2^M+1}P_c$
DPIM	$\frac{2^M+3}{2}$	$\frac{2^M+3}{2M}R_b$	$\frac{2}{2^M+3}P_c$
DHPIM	$\frac{2^{M-1}+2\alpha+1}{2}$	$\frac{2^{M-1}+2\alpha+1}{\alpha M}R_b$	$\frac{3\alpha}{\left(2^M+4\alpha+2\right)}P_c$
DPPIM	$2^{M-1}+\alpha$	$\frac{2^M+2\alpha}{\alpha M}R_b$	$\frac{4+3\alpha}{2^{M+1}+4\alpha}P_c$
DAPIM	$\frac{2^{M-1}+3}{2}$	$\frac{2^{M-1}+3}{2M}R_b$	$\frac{1+\beta}{2^{M-1}+3}P_c$
FDPIM	2^M+4	$\frac{2^M+4}{M}R_b$	$\frac{3}{2^M+4}P_t$
FDAPIM	2^M+3	$\frac{2^M+3}{M}R_b$	$\frac{1+\beta}{2^M+3}P_c$
SPPM	$1+2^{M-1}$	$\frac{2^{M-1}+1}{M}R_b$	$\frac{3}{2^M+2}P_c$
SDPPM	n_s	n_sR_b/M	$2P_c/n_s$

the latter represents the transmitted information using the number of space slots between adjacent pulses.

Other modulation methods are developed by changing the symbol length and the amplitude, width, and number of pulses, based on these two modulation methods. An analysis reveals that the bandwidth requirements of different modulation modes increase with an increase in their average symbol length and decrease with an increase in the pulse width [59].

Slot Error Rate Analysis

In an optical wireless communication system, the information sent by a signal source is encoded and modulated into a group of information sequences composed of multiple time slots. Each time slot takes the value of "1" or "0," according to the presence or absence of an optical pulse, respectively. It is then loaded into the laser and transmitted through an optical antenna. Because a laser transmission in the atmosphere is affected by turbulence, background-light noise, path loss, and other factors, the information will be distorted or lost, and the receiver will receive the wrong information. The slot error rate (SER) refers to the probability of a slot error in

each bit of symbol information, that is, the sum of the probability of sending "1" and receiving "0" and the probability of sending "0" and receiving "1." The slot error rate performance of various modulation methods is analyzed below [50].

Slot Error Rate Analysis Under a Gaussian Channel Model

Figure 3.8 shows the slot error rate comparison of different modulation modes under different modulation bit numbers, without considering the influence of turbulence. The pulse-width parameter was 1 and the amplitude parameter was 2. It can be seen from the three figures that the slot error rate of OOK modulation is independent of the modulation order M; the higher the signal-to-noise ratio (SNR), the smaller the slot error rate for each modulation mode. When the signal-to-noise ratio is constant, the larger the M value, the smaller the slot error rate. When the M value is fixed, the slot error rate of OOK is the largest, and the slot error rate of PPM is the smallest. Moreover, with the increase in M, the gap between the slot error rates of various modulation methods increases [1].

Average Slot Error Rate Under a Weak Turbulence Channel Model

A comparison of the average time-slot error rates of different modulation modes in weak turbulence is shown in Fig. 3.9. The three subgraphs compare the average time-slot error rates of different modulation modes for different transmission distances and different modulation bits in weak turbulence.

Average Slot-Error Rate Under a Medium–Strong Turbulence Channel Model

Figure 3.10 shows a comparison of the slot error rate of each modulation mode under different transmission distances and modulation bit numbers in the case of medium–strong turbulence. It can be seen from the three subgraphs that the change trend of various modulation modes under the corresponding conditions is predominately the same as that under weak turbulence. However, the difference between dual-amplitude pulse-interval modulation (DAPIM), 2MPPM, and separated double-pulse-position modulation (SDPPM) under moderate–strong turbulence is larger than that under weak turbulence, and the average slot error rate under moderate–strong turbulence is higher [53].

Average Slot-Error Rate Under a Strong Turbulence Channel Model

Figure 3.11 shows the average slot error rate comparison of different modulation modes with different a values and different modulation bit numbers under strong turbulence. It can be seen from the three subgraphs that, with the change in the values of a and M, the order of magnitude of the above modulation modes does not change. When the value of a is fixed, the average slot error rate of each modulation mode decreases with an increase in the number of modulation bits, M. When the value of M is fixed, the average slot error rate of each modulation mode decreases with increasing a. The average slot error rate in the three graphs is 10^{-4}–10^{-5}, which is higher than the average slot error rate in the medium–strong turbulence.

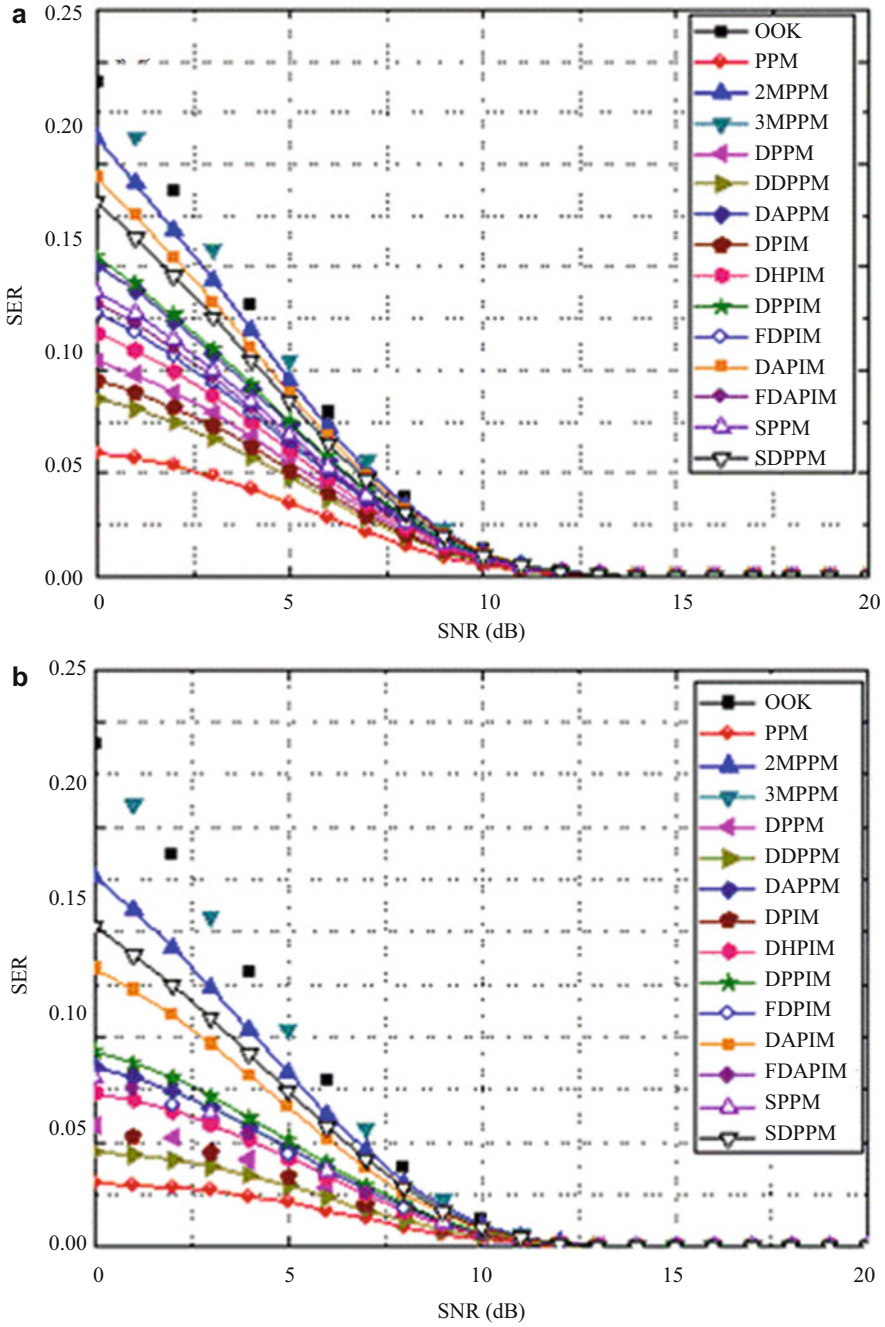


Fig. 3.8 Slot error rate of each modulation mode in a Gaussian channel [59]. (**a**) $M = 4$ bits; (**b**) $M = 5$ bits; and (**c**) $M = 6$ bits

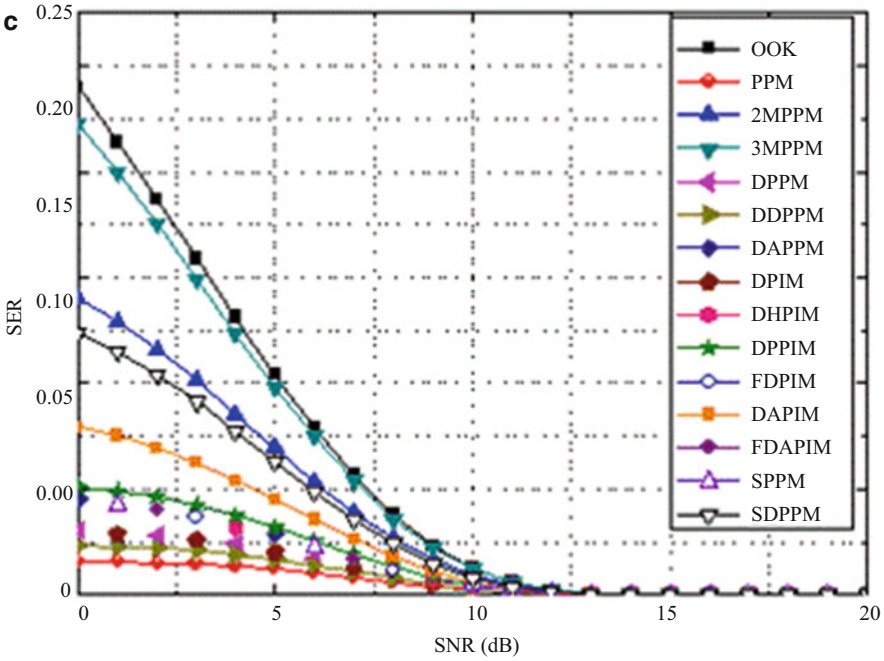


Fig. 3.8 (continued)

This study analyzes the performance of the average slot error rate of different modulation modes when the light intensity obeys lognormal, gamma–gamma, and K distributions under the three turbulence conditions. It obtains the calculation formula of the average slot error rate of each modulation mode through mathematical derivation and carries out numerical simulations on these formulas.

From the analysis process and numerical simulation results, it can be observed that the slot error rate of the OOK modulation mode is independent of the modulation order M; the higher the signal-to-noise ratio, the smaller the slot error rate for each modulation mode. When the signal-to-noise ratio is fixed, the slot error rate decreases with an increase in M, and the change trend of the average slot error rate under the three turbulence conditions is the same. The slot error rate of OOK is the largest and that of PPM is the smallest. The average slot error rate increases with an increase in the turbulence intensity.

Using the same modulation method in the same turbulence channel, the average slot error rate performance of a modulation method can effectively be improved by increasing the modulation ratio M, reducing the information-transmission distance L or increasing the channel parameter α [50].

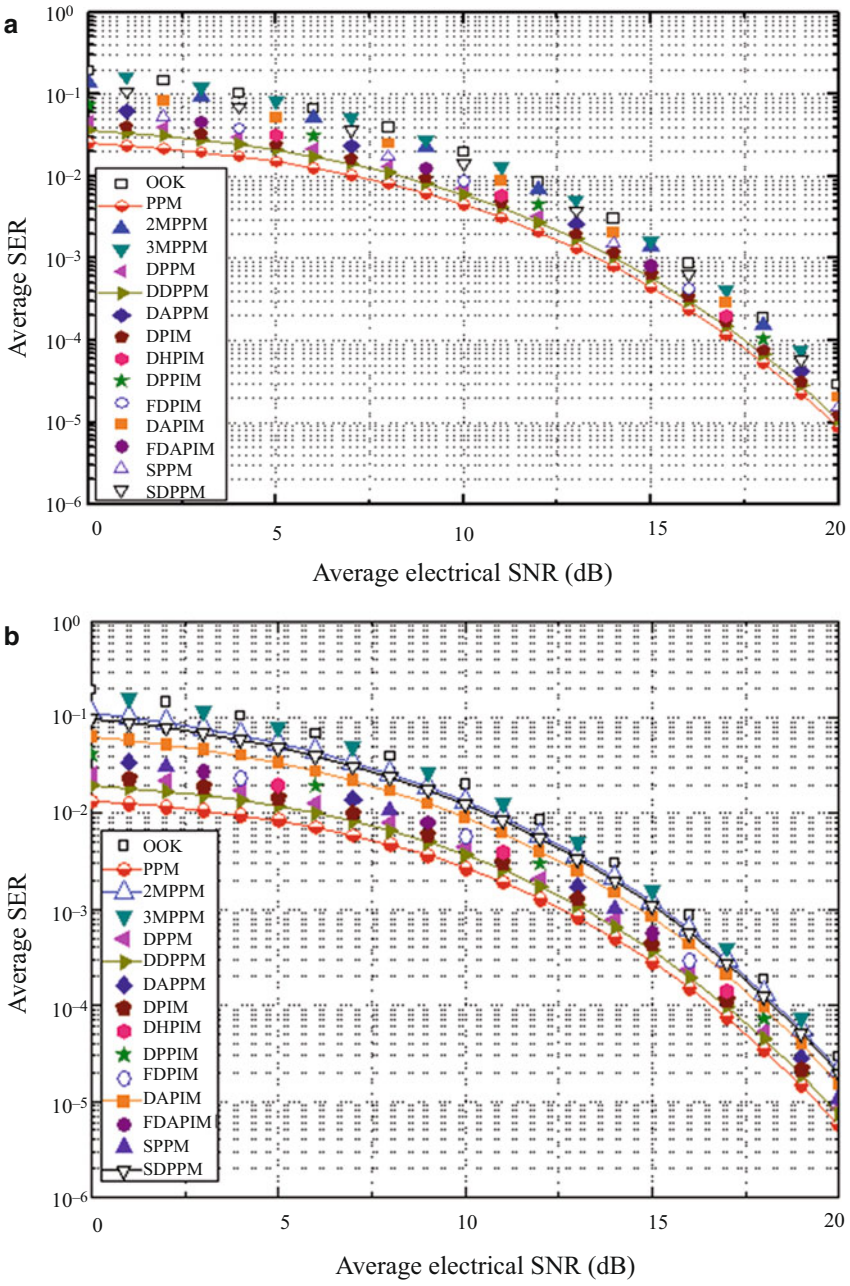


Fig. 3.9 Average slot error rate of each modulation mode in a weak turbulence channel [59]. (**a**) $L = 5000$ m, $M = 5$ bits, $C_n^2 = 1 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$; (**b**) $L = 5000$ m, $M = 6$ bits, $C_n^2 = 1 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$; (**c**) $L = 3000$ m, $M = 6$ bits, $C_n^2 = 1 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$

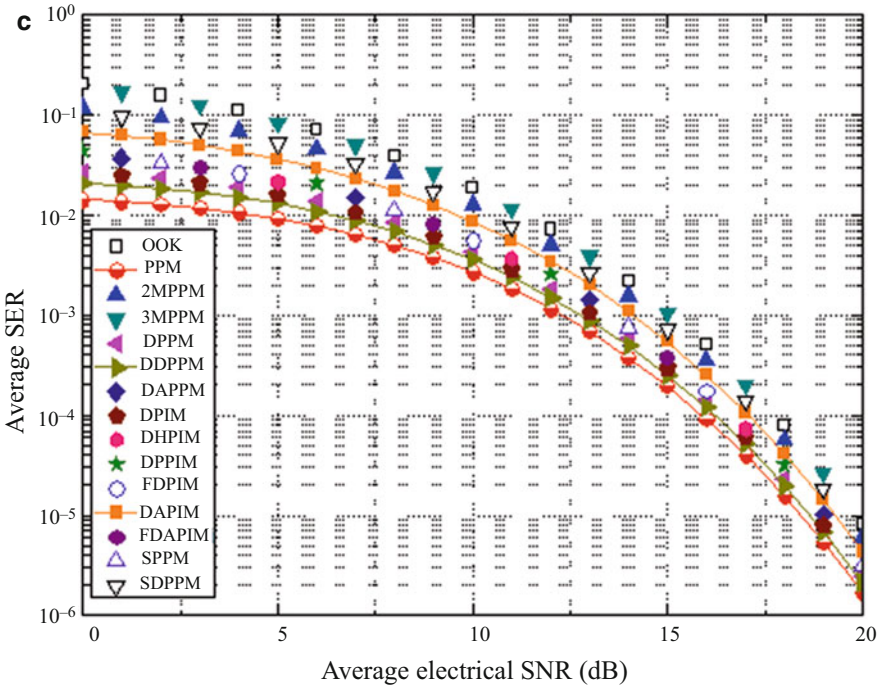


Fig. 3.9 (continued)

Channel Capacity

The channel capacity refers to the maximum average information rate that a channel can transmit. Next, the transmission capacity of various modulation modes and the average channel capacity of different modulation modes in turbulent channels are analyzed.

Transmission Capacity

Figure 3.12 shows a comparison of the transmission capacity of different modulation modes. It can be observed from the figure that the transmission capacity of each modulation mode increases logarithmically with an increase in the number of modulation bits. The transmission capacity of DHPIM is the largest, approaching 4 bit/s, followed by PPM and DPIM. The transmission capacity of these two modulation modes approaches 2 bit/s, and the smallest is that of PPM. For the modulation mode with variable pulse-width parameters, the transmission capacity decreases with an increase in the pulse-width parameter.

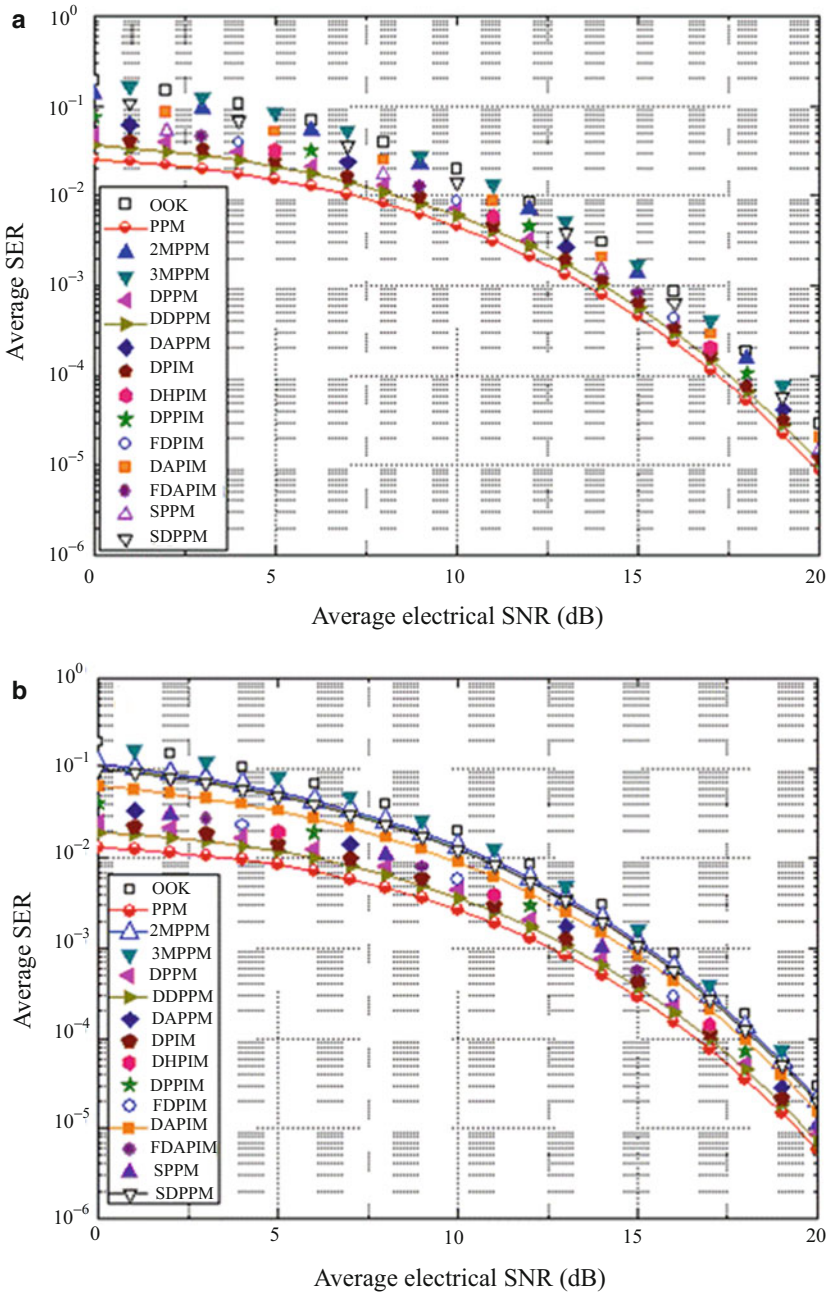


Fig. 3.10 Average slot error rate of each modulation mode in a medium–strong turbulence channel [59]. (**a**) $L = 5000$ m, $M = 5$ bits, $C_n^2 = 9 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$; (**b**) $L = 5000$ m, $M = 6$ bits, $C_n^2 = 9 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$; (**c**) $L = 3000$ m, $M = 6$ bits, $C_n^2 = 9 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$

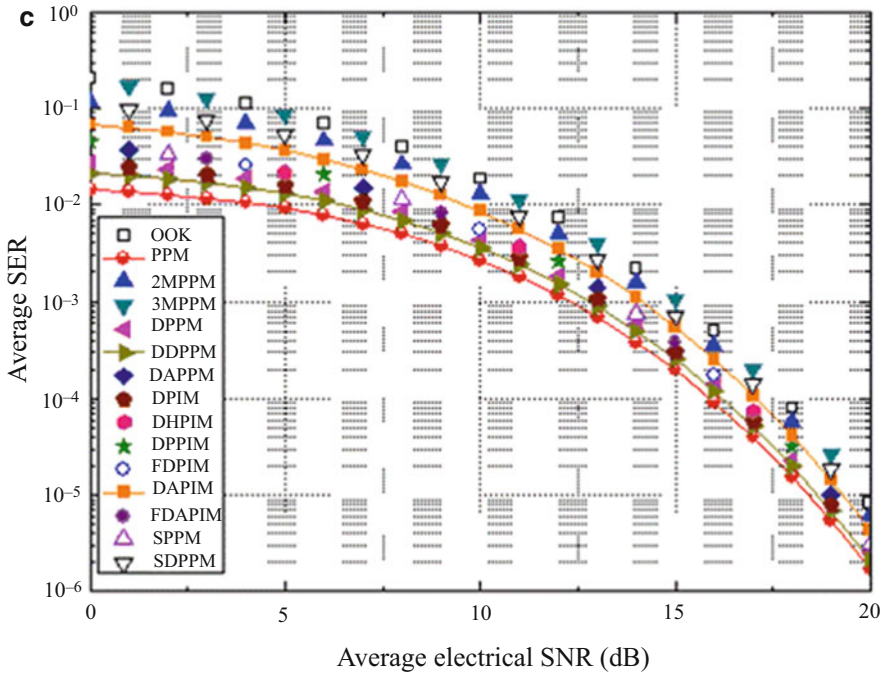


Fig. 3.10 (continued)

Average Channel Capacity Under a Turbulent Channel

Figure 3.13 shows a comparison of the average channel capacity at different transmission distances and light wavelengths in a turbulent channel. The medium–strong turbulence channel was simulated using the gamma–gamma distribution model and distinguished by different atmospheric refractive-index constants.

Figure 3.14 shows a comparison of the average channel capacity when the light intensity follows the K distribution under strong turbulence. It can be observed from the figure that the average channel capacity increases with an increase in the signal-to-noise ratio and channel parameter.

Average Channel Capacity of Different Modulation Modes and Turbulence Intensities

Figure 3.15 shows a comparison of the average channel capacity of different modulation methods at different transmission distances and modulation bits in weak turbulence. It can be observed from the figure that the average channel capacity increases with an increase in the peak transmission power, decreases with an increase in the transmission distance, and decreases with an increase in the number of modulation bits. The average channel capacity of OOK was the largest and that of PPM was the smallest [48].

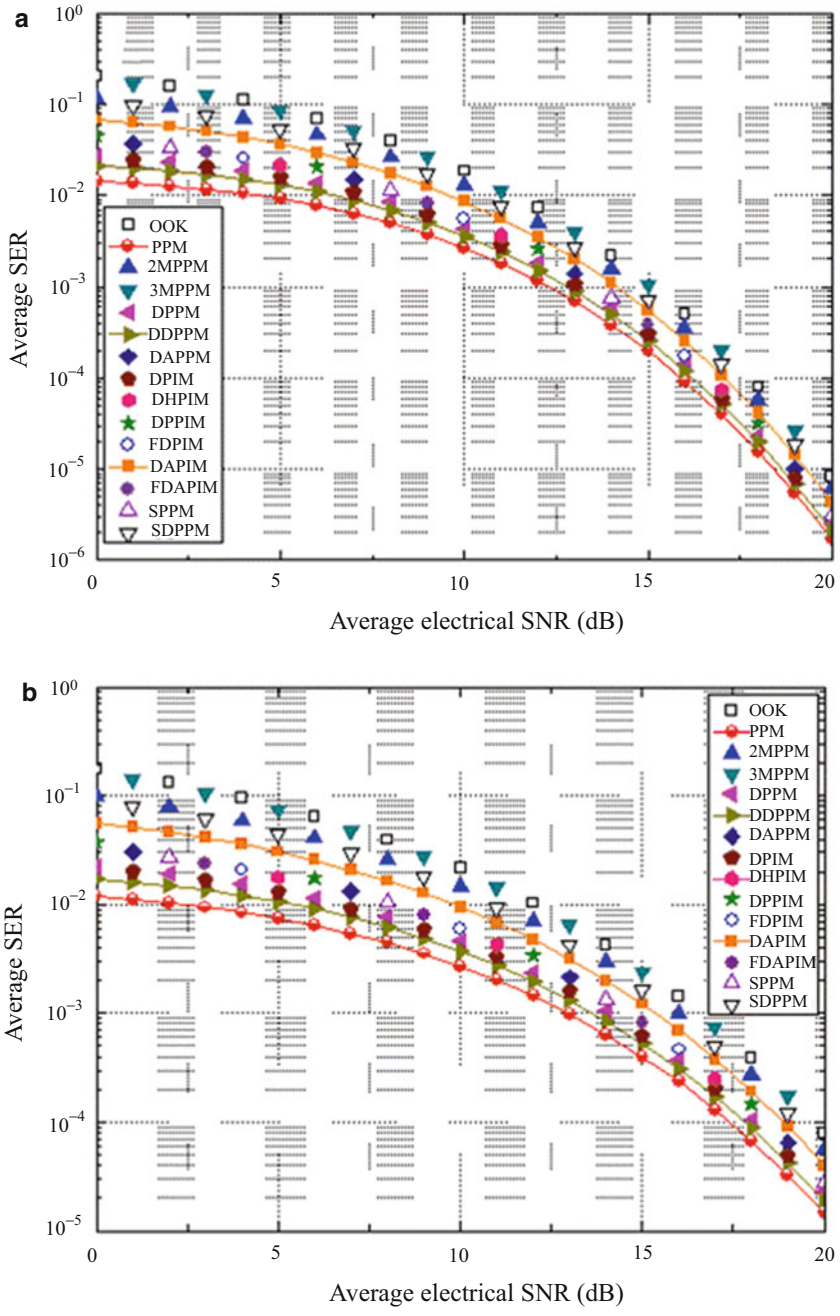


Fig. 3.11 Average slot error rate of each modulation mode under a strong turbulence channel [59]. (**a**) $a = 4$, $M = 5$ bits; (**b**) $a = 4$, $M = 6$ bits; (**c**) $a = 2$, $M = 6$ bits

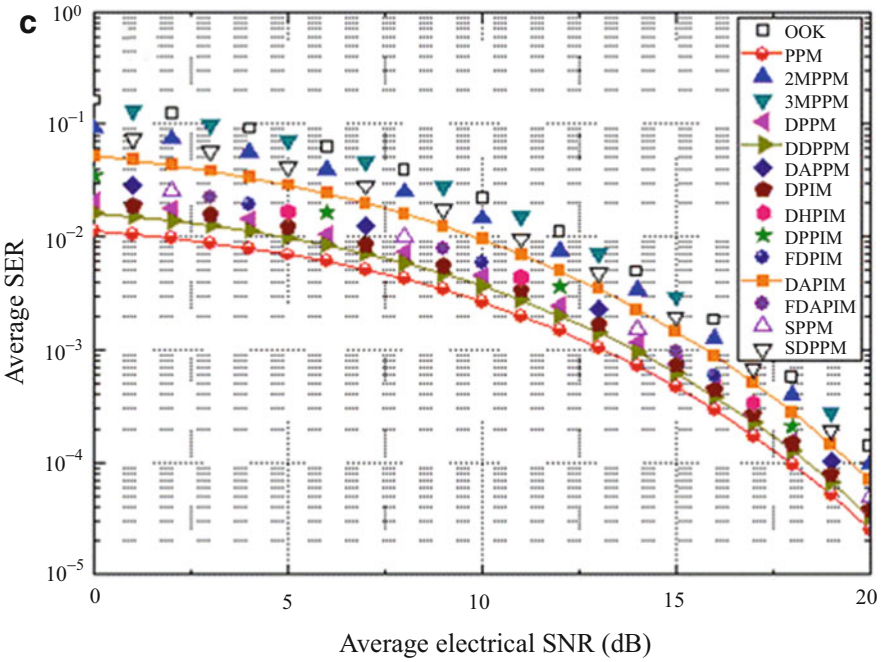


Fig. 3.11 (continued)

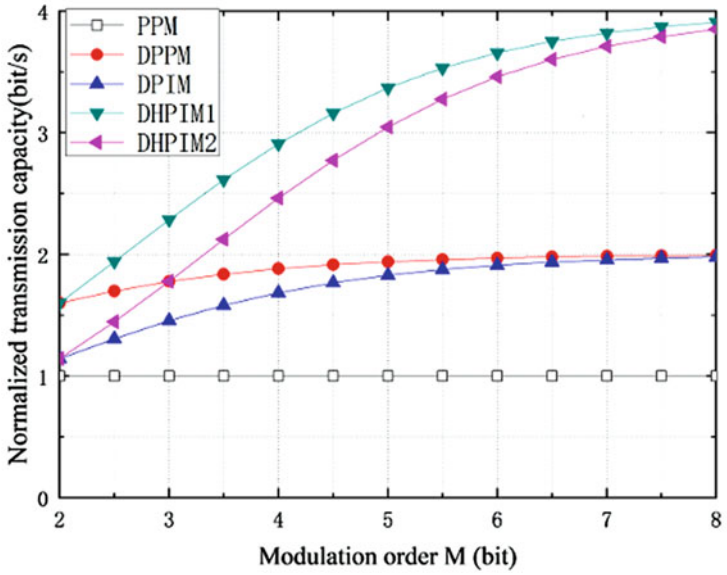


Fig. 3.12 Normalized transmission capacity of various modulation modes [51]

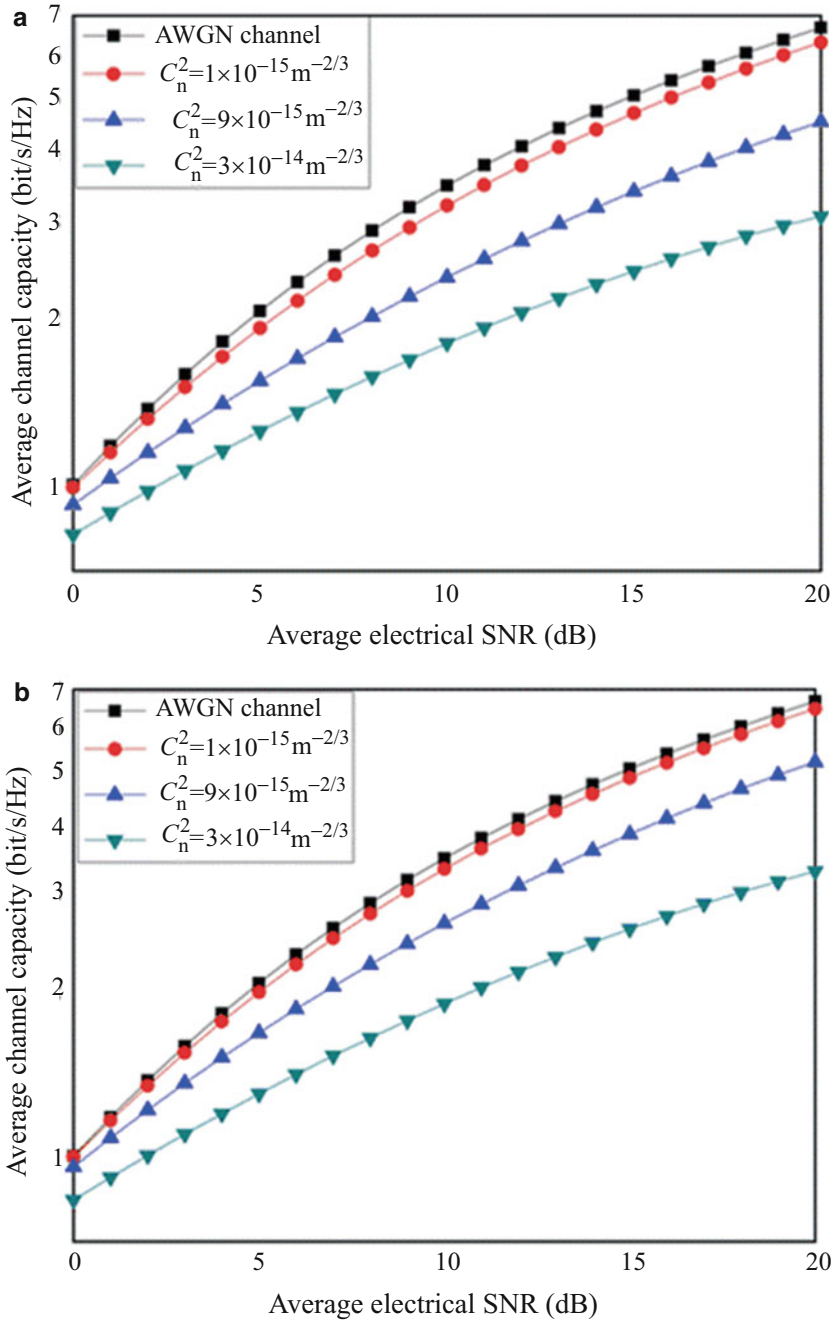


Fig. 3.13 Average channel capacity under a Gaussian turbulence channel [51]. (**a**) $L = 3000$ m, $\lambda = 850$ nm; (**b**) $L = 3000$ m, $\lambda = 1550$ nm; (**c**) $L = 5000$ m, $\lambda = 850$ nm

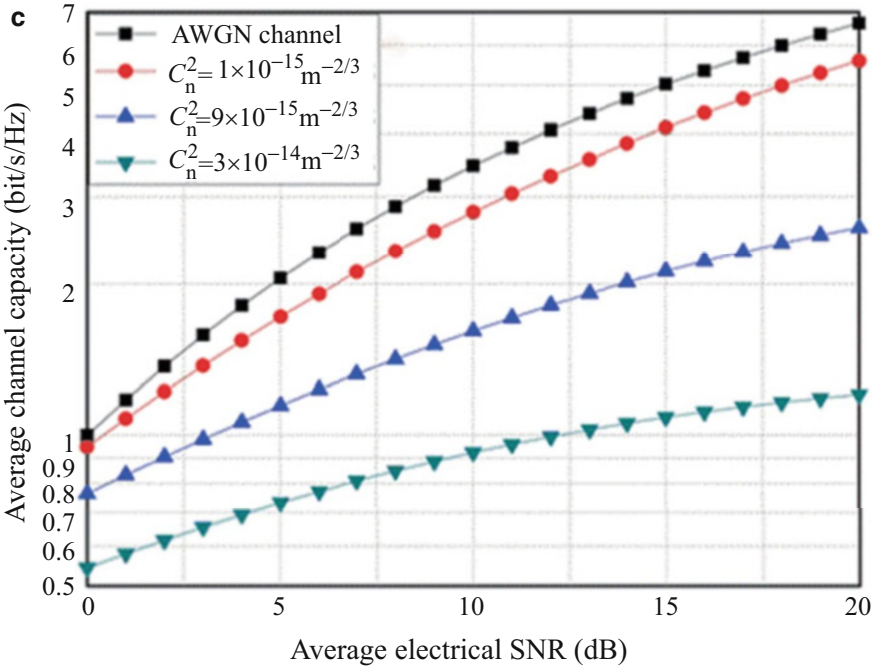


Fig. 3.13 (continued)

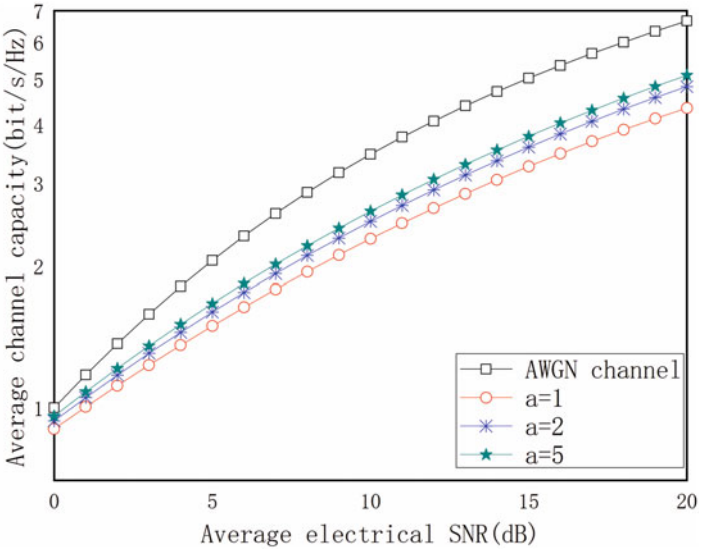


Fig. 3.14 Average channel capacity under a K distribution in a strong turbulence channel [61]

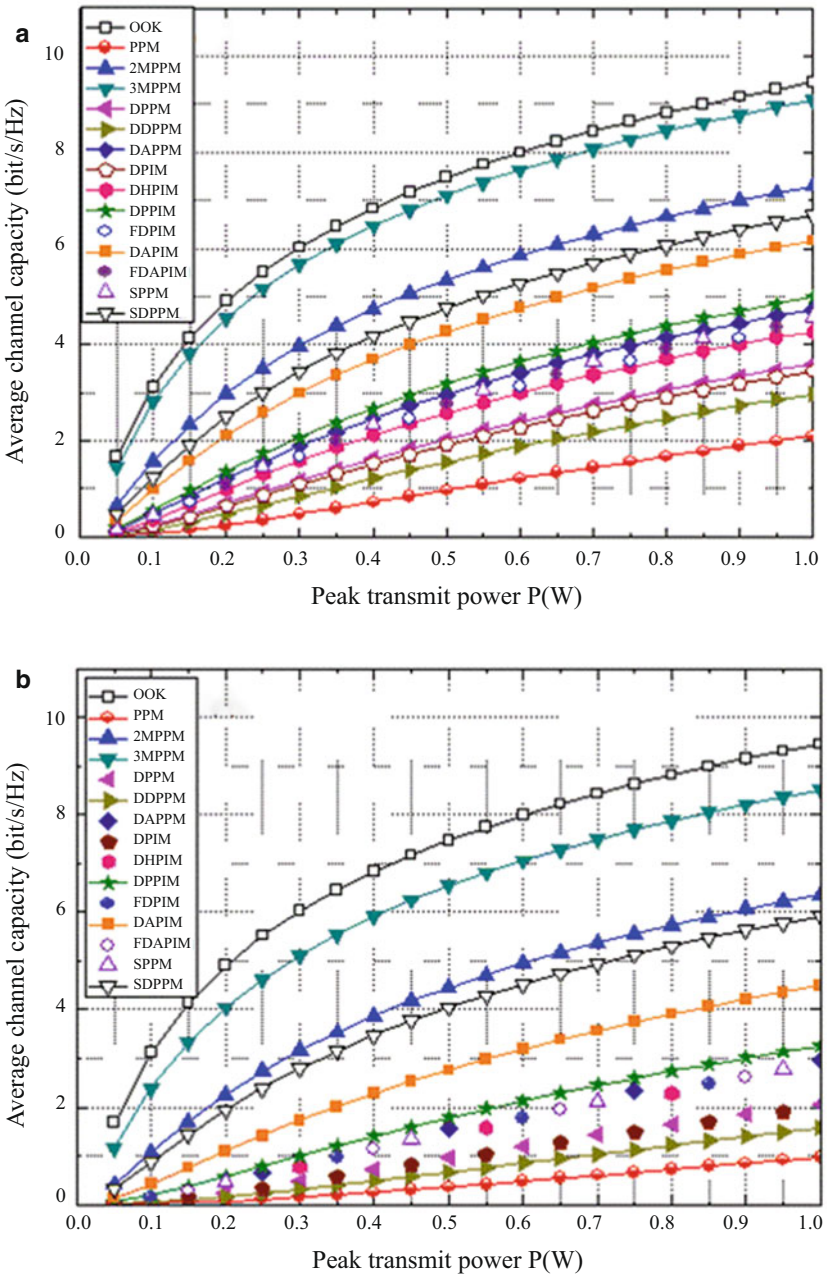


Fig. 3.15 Average channel capacity of each modulation mode under a weak turbulence channel [73]. (**a**) $L = 5000$ m, $M = 5$ bits, $C_n^2 = 1 \times 10^{-15} m^{\frac{-2}{3}}$; (**b**) $L = 5000$ m, $M = 6$ bits, $C_n^2 = 1 \times 10^{-15} m^{\frac{-2}{3}}$; (**c**) $L = 3000$ m, $M = 6$ bits, $C_n^2 = 1 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$

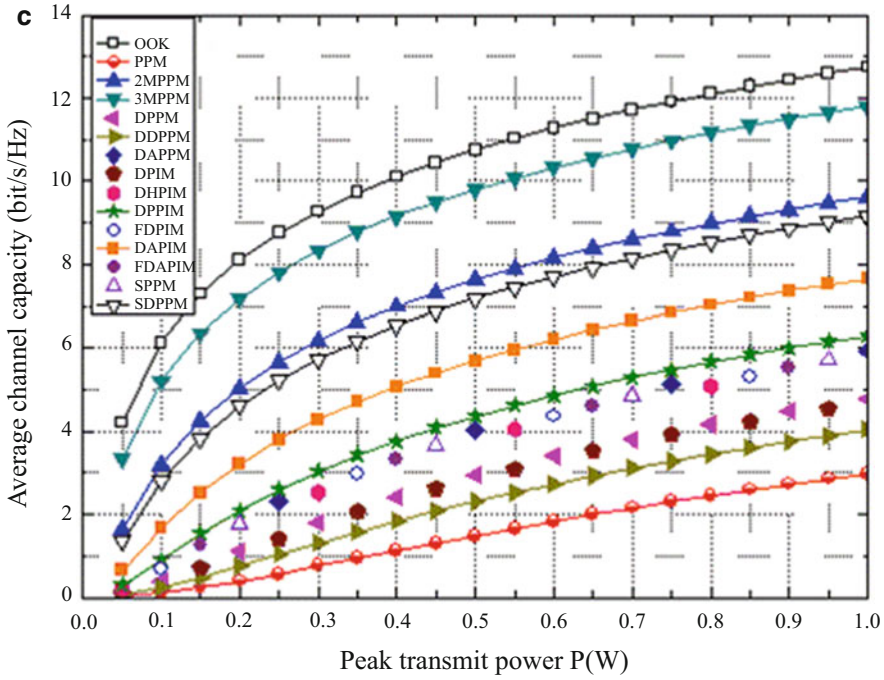


Fig. 3.15 (continued)

Figure 3.16 shows a comparison of the average channel capacity of various modulation methods at different transmission distances and modulation bits in moderate–strong turbulence. From these three subgraphs, we can see that the change trend of the average channel capacity under the same conditions is consistent with that under weak turbulence.

Figure 3.17 shows a comparison of the average channel capacity of each modulation mode under different transmission distances, modulation bits, and α values in strong turbulence. From the four sub-figures (a, b, c, and d), it can be seen that the influence of the transmission distance and modulation bit number on the average channel capacity of each modulation mode is the same as that of weak and medium–strong turbulence. The larger the channel parameter a value, the larger the average channel capacity; however, the difference is not obvious. Compared with the medium–strong turbulence, the average channel capacity of strong turbulence under the same conditions is approximately 1 bit/s/Hz lower.

This section analyzes and obtains the average channel capacity of each modulation mode under three turbulence channels, namely weak, medium, and strong. The simulation results show that the average channel capacity of OOK modulation is the largest and independent of the number of transmitted information bits. It is followed

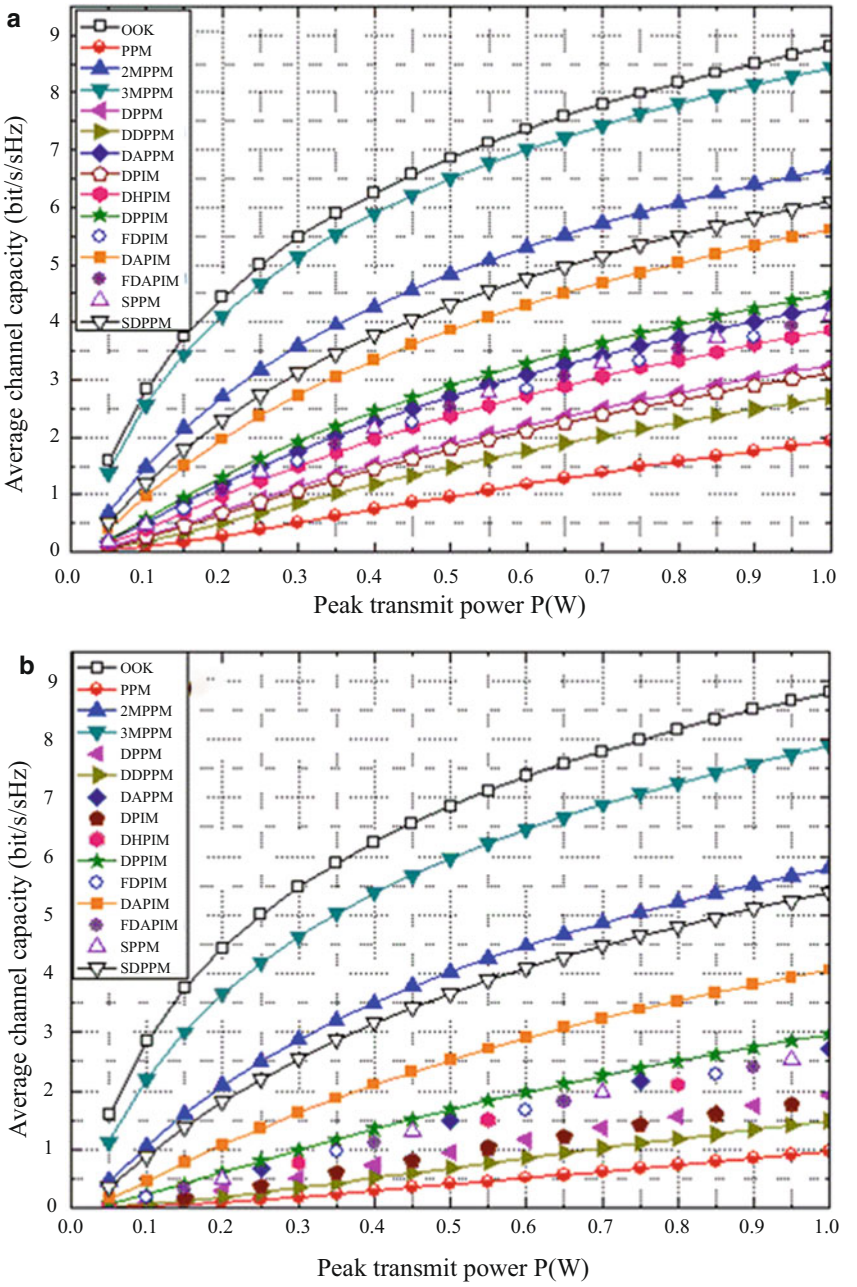


Fig. 3.16 Average channel capacity of each modulation mode under a medium–strong turbulence channel [73]. (**a**) $L = 5000$ m, $M = 5$ bits, $C_n^2 = 9 \times 10^{-15} m^{\frac{-2}{3}}$; (**b**) $L = 5000$ m, $M = 6$ bits, $C_n^2 = 9 \times 10^{-15} \mathrm{m}^{\frac{-2}{3}}$; (**c**) $L = 3000$ m, $M = 5$ bits, $C_n^2 = 9 \times 10^{-15}$ $\mathrm{m}^{\frac{-2}{3}}$

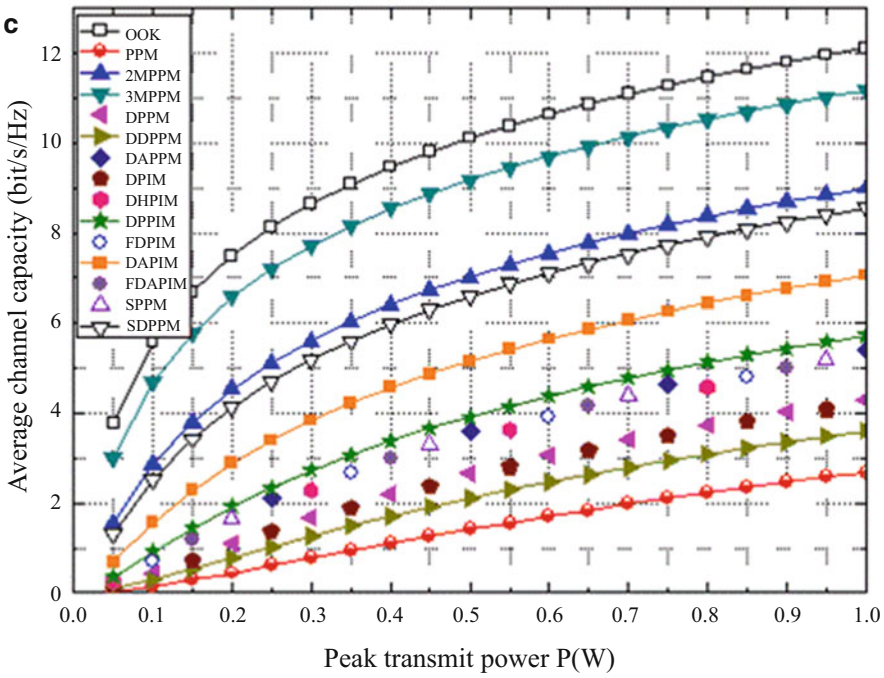


Fig. 3.16 (continued)

by 3MPPM, 2MPPM, SDPPM, then DAPIM, DPPIM, DAPPM, SPPM, FDAPIM, FDPIM, DHPIM, DPPM, and DPIM, and finally, dual-duration pulse-position modulation (DDPPM) and PPM.

When the transmission distance and modulation bit number are fixed, the stronger the turbulence intensity, the smaller the average channel capacity. The average channel capacity in weak turbulence is close to that in medium–strong turbulence; the former is approximately 0.5 bit/s/Hz higher than the latter, and the average channel capacity in medium–strong turbulence is approximately 1 bit/s/Hz higher than that in strong turbulence. The larger the transmission distance, the smaller the average channel capacity.

The average channel capacity at a transmission distance of 3000 m is approximately 3 bit/s/Hz higher than that at 5000 m; the larger the number of information bits transmitted, the smaller the average channel capacity. The average channel capacity with four modulation bits is (1–2) bit/s/Hz higher than when the number of modulation bits is five. Therefore, in terms of the average channel capacity, the best modulation method is MPPM. The average channel capacity can be improved by reducing the transmission distance and increasing the peak transmission power.

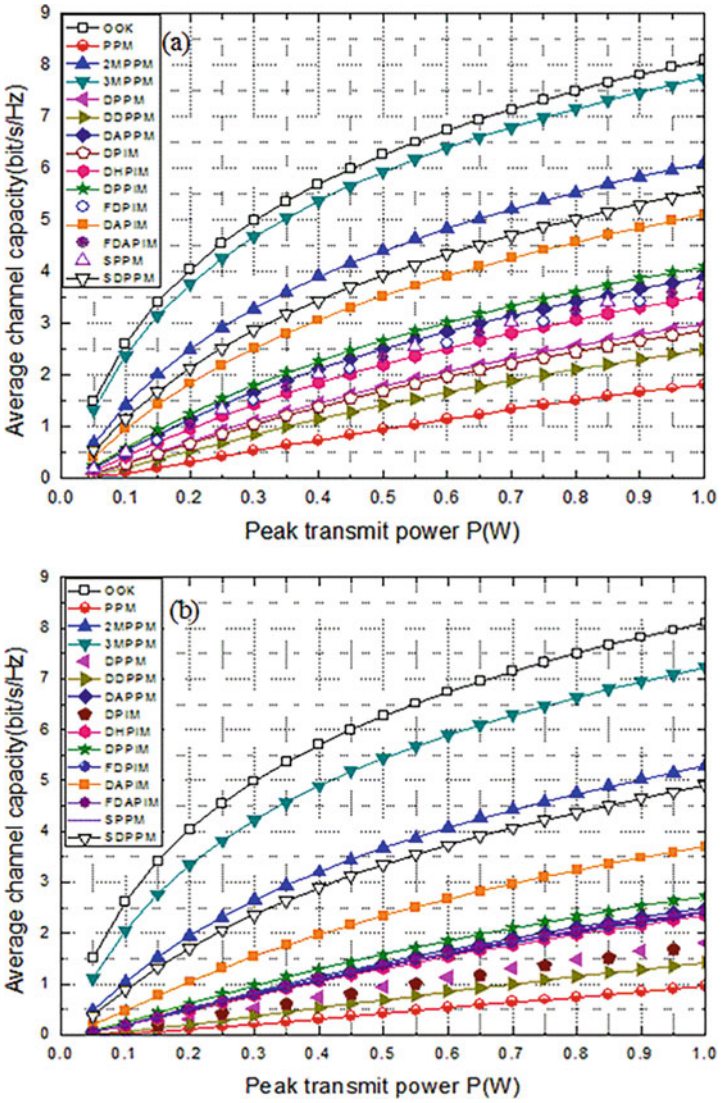


Fig. 3.17 Average channel capacity of each modulation mode under strong turbulence [59]. (**a**) $L = 5000$ m, $M = 5$ bits, $a = 4$; (**b**) $L = 5000$ m, $M = 6$ bits, $a = 4$; (**c**) $L = 3000$ m, $M = 6$ bits, $a = 4$; (**d**) $L = 3000$ m, $M = 6$ bits, $a = 2$

Adaptive Threshold Detection

Figure 3.18 shows the simulation results of the change in bit-error rate with the signal-to-noise ratio. The dotted line in the figure represents the fixed threshold-detection method, and the solid line represents the adaptive threshold-detection method when the sampling numbers M are 8, 16, 32, and 64.

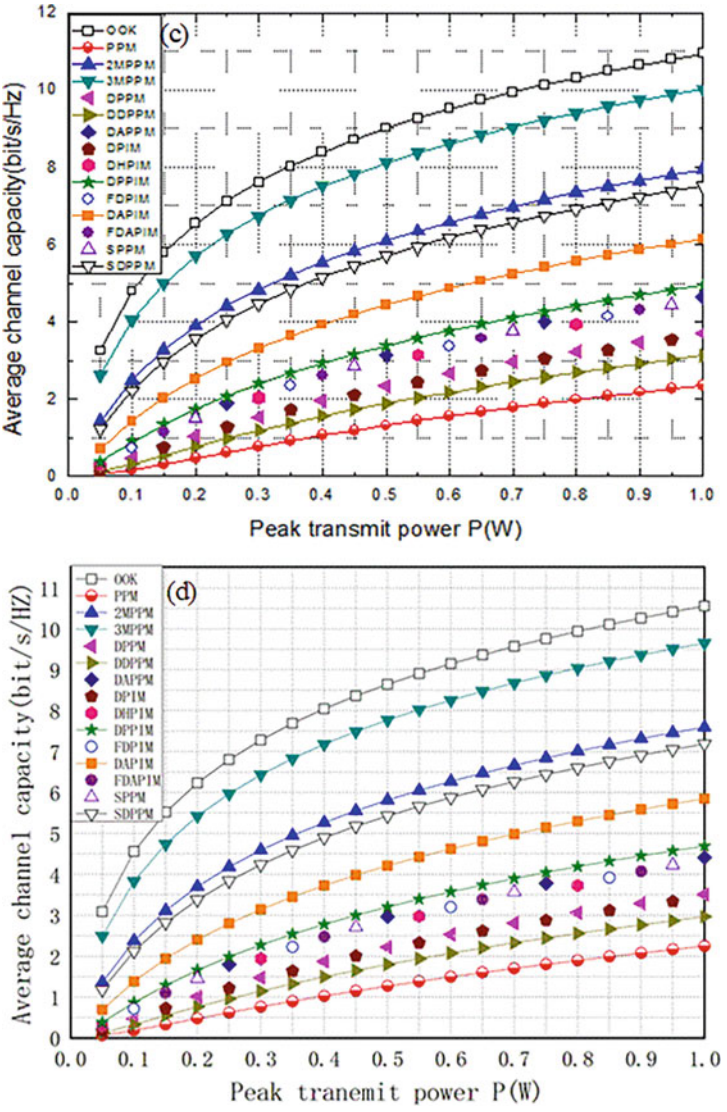


Fig. 3.17 (continued)

Figure 3.18 shows the change curve of the BER with the SNR under the fixed and adaptive threshold-detection methods. It can be observed that the bit-error rate of adaptive threshold detection is significantly lower than that of the fixed threshold-detection method at the same signal-to-noise ratio. Figure 3.19 shows the relationship curve between the signal-to-noise ratio and bit-error rate when 8-PPM, 16-PPM, and 256-PPM signals are used for transmission and three-aperture reception is used.

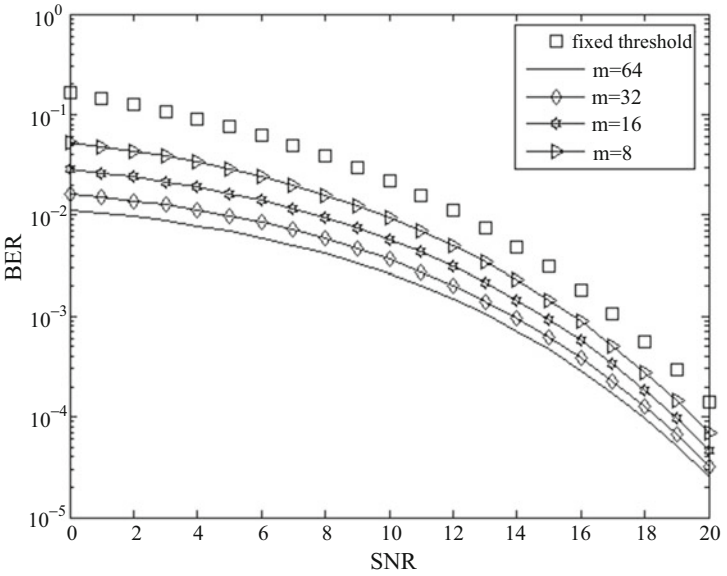


Fig. 3.18 Change curve of BER with SNR [61]

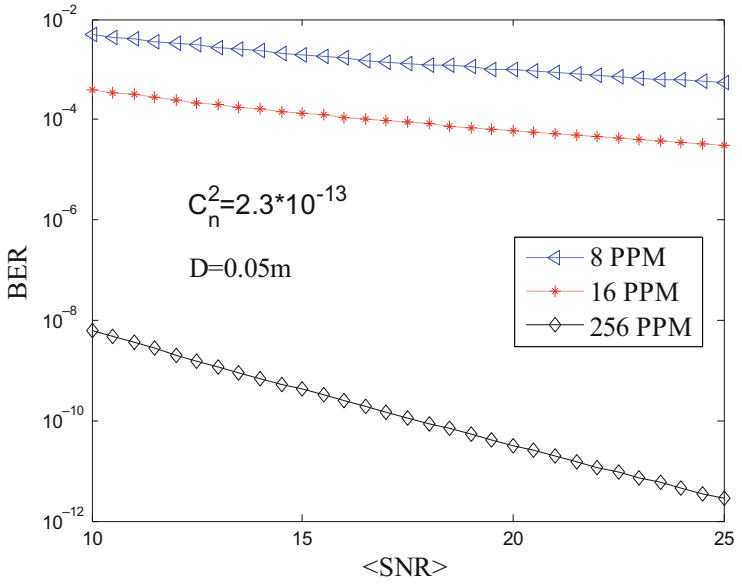


Fig. 3.19 BER curve of receiver [62]

It can be observed that the bit-error rate of the received signal decreases significantly when the number of PPM-signal time slots increases.

Figures 3.20, 3.21, and 3.22 show the change curves of the bit-error rate with the signal-to-noise ratio when the sampling numbers of the measured data are 8, 16,

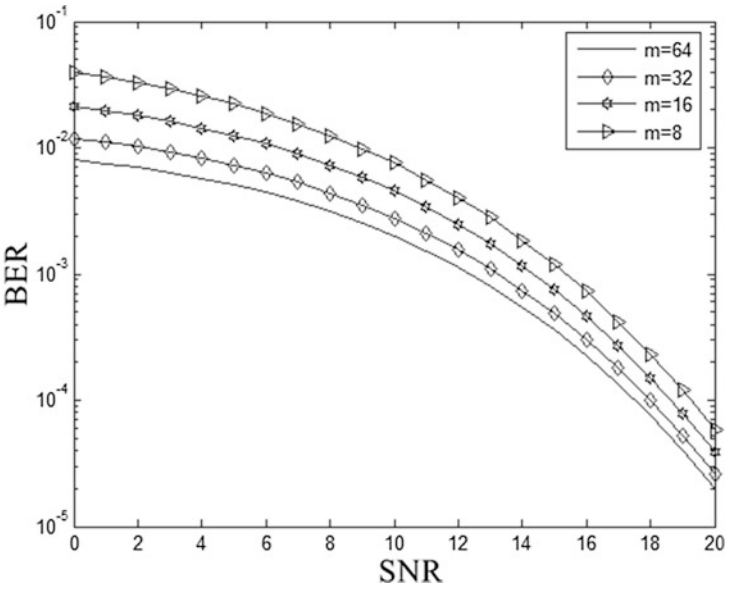


Fig. 3.20 Change curve of the bit-error rate with SNR in heavy rain [61]

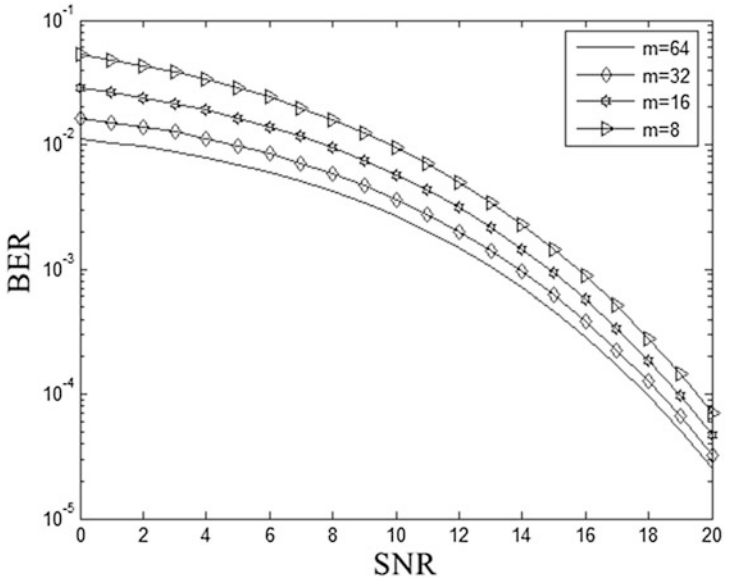


Fig. 3.21 Change curve of the bit-error rate with SNR in moderate rain [61]

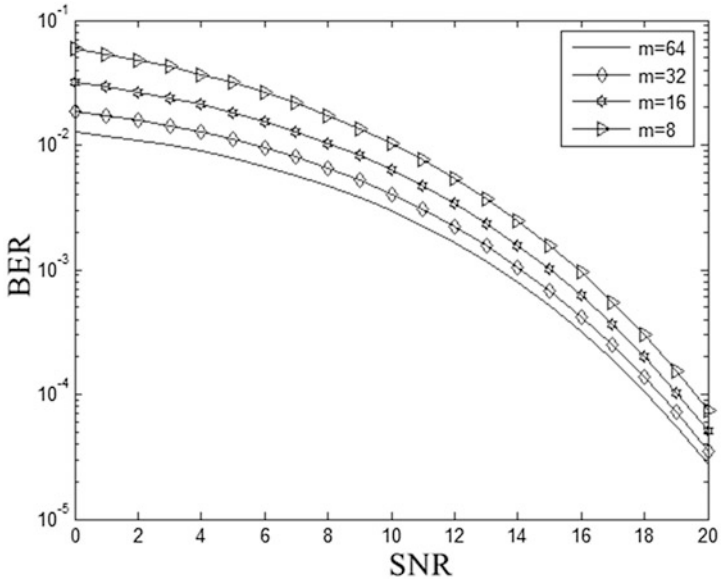


Fig. 3.22 Change curve of the bit-error rate with SNR in light rain [61]

32, and 64 in heavy, moderate, and light rain, respectively. When the number of measured data samples increases from 8 to 64, it can be seen from the figure that, under the same signal-to-noise ratio, the larger the number of samples, the lower the system error rate, and the more obvious the performance improvement of the adaptive detection method. When the number of samples exceeds 32, the performance improvement tends to flatten.

Figure 3.23 shows the change curve of the received-signal error rate with SNR in heavy, moderate, and light rain with 64 samples. It can be seen that when the SNR is the same, with the same number of symbol samples, the bit-error rate of the received signal in heavy rain is smaller than that in moderate and light rain. The received signal quality obviously improves, and the bit-error rate in moderate and light rain increases.

Factors Affecting the Laser-Modulation Rate

A high modulation rate is an important index for optical wireless communications. The main factors affecting optical wireless modulation were analyzed. It can be seen from Fig. 3.24 that when the bias current I_0 changes from 40 mA to 70 mA, the resonance peak of the relaxation oscillation shifts to the right, and the relaxation oscillation frequency f_0 increases from 6.1 GHz to 13.5 GHz.

It can be seen that the relaxation oscillation frequency f_0 can be increased by increasing the bias current I_0. Therefore, to improve the laser-modulation frequency, we can increase the steady-state photon density S_0 by increasing the bias current I_0 and then increase the relaxation oscillation frequency f_0. As shown in Fig. 3.25,

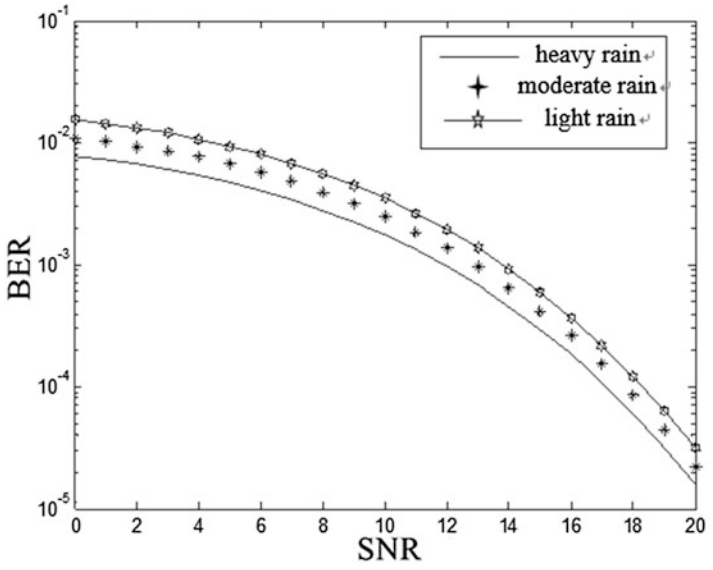


Fig. 3.23 Change curve of the bit-error rate with SNR in different weather conditions [61]

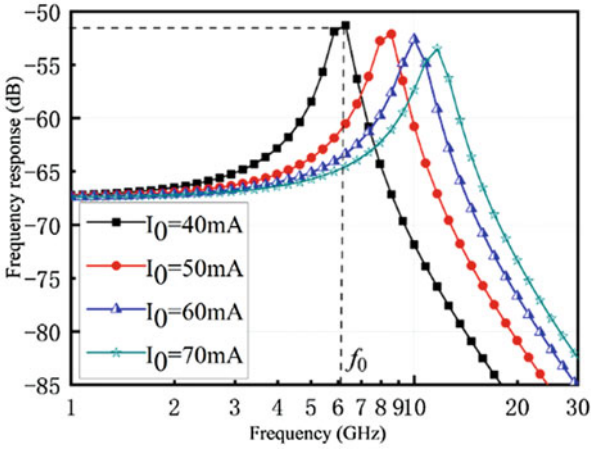


Fig. 3.24 Frequency-response curve under different bias currents I_0 [63]

when the bias current I_0 changes from 40 mA to 70 mA, the output optical power P increases from 14.5 mW to 27 mW.

Increasing the bias current I_0 can increase the relaxation oscillation frequency and reduce the relaxation oscillation amplitude, thereby shortening the relaxation oscillation process. Therefore, pulse waveform distortion can be suppressed by increasing

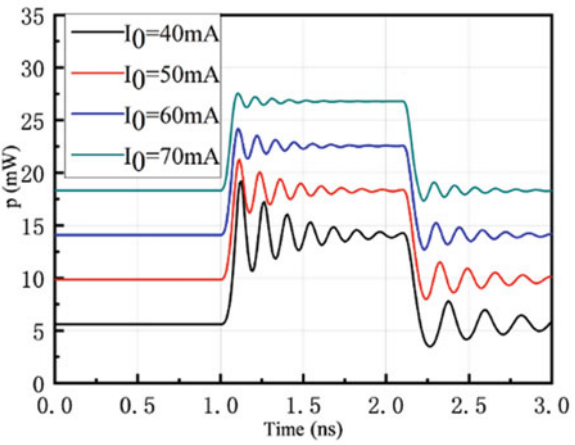


Fig. 3.25 Pulse-response curve under different bias currents I_0 [63]

the steady-state photon density S_0 (bias current I_0). It is difficult to increase the modulation frequency of a laser by increasing the current bias.

Summary

Pulse-position modulation includes PPM, DPPM, and other related and evolved modulation methods. The advantages of pulse-like position modulation are that the modulation modes are simple and the hardware circuits are convenient; however, the influence of the channel on its communication performance cannot be ignored. Researchers are investigating optical wireless communication in the following aspects.

1. Multiple transmission/reception technologies have been adopted to increase the communication capacity and suppress channel noise [64]. PAM4 is a type of PAM technology. The PAM signal is a popular signal-transmission technology after non-return-to-zero (NRZ) and is also representative of multistep modulation technology. It has been widely used in the high-speed signal interconnection field.
2. A longer-wavelength laser is used to transmit signals to avoid the impact of various weather factors, such as rain, fog, and sandstorms, on communication performance [65].
3. New modulation methods, such as reverse modulation [66], have been adopted to meet the needs of specific applications.
4. New coding technologies, such as polarization codes [72], have been investigated, and error correction codes used to suppress channel-noise interference, such as atmospheric turbulence, and improve the anti-interference ability of optical wireless communication.

5. New optical multiple-in-multiple-out (OMIMO) technologies have been studied, such as optical spatial modulation (OSM) [67], and the spatial laser index number has been used to carry additional information, effectively improving the transmission rate and energy efficiency of the system.
6. Coherent detection technology has been investigated to compensate for wavefront distortion caused by atmospheric turbulence and improve the ability to suppress it [68, 69].
7. The influence mechanism of atmospheric turbulence on optical signal transmissions has been explored and the optical communication coding and modulation modes have been adaptively changed, according to the intensity of the atmospheric turbulence, to approach the channel capacity limit [70, 71].

Through analysis and comparison, it can be seen that OOK modulation mode is simple and has low bandwidth requirements; however, the average transmission power is high, the power utilization rate is low, and the anti-interference ability is poor. PPM mode is simple, its power utilization is the highest, the average transmission power is small, and its anti-interference ability is strong; however, the bandwidth efficiency is poor, and simultaneous symbol and frame synchronization are required.

DPIM has a complex modulation structure, shortened symbol length, high bandwidth efficiency, and a higher power efficiency than OOK, without the symbol synchronization. DHPIM mode has a complex structure, which further shortens the symbol length, and improves the bandwidth efficiency and transmission capacity. It does not need symbol synchronization, which greatly simplifies the complexity of the system implementation, and it can use higher modulation modes in exchange for higher power utilization.

Pulse-like position modulation is a type of pulse modulation whose bandwidth occupation is greater than that of sine modulation. This is a disadvantage that is still difficult to overcome at present. The continuous solutions of key technologies in optical wireless communication have created opportunities for its industrialization and application. The continuous improvement of this technology will certainly promote the development and application of optical wireless communication theory and technology in China [73, 74].

References

1. Ke X Z, Yin Z Y. Coding theory in wireless laser communication system [M]. Beijing: Science Press, 2009: 8–11.
2. Ke X Z. Optical Wireless Communication [M]. Beijing: Science Press, 2016: 10–14.
3. J.R. Pierce. Optical Channels: Practical Limits with Photon Counting [J]. Communications, 1978, 26(12):1819–1821.
4. Prati G, Gagliardi R. Block encoding and decoding for the optical PPM channel(Corresp.) [J]. IEEE Transactions on Information Theory, 1982, 28(1):100–105.
5. Prati G, Gagliardi R. Decoding with Stretched Pulses in Laser PPM Communications [J]. IEEE Transactions on Communications, 1983, 31(9):1037–1045.

6. Garrett, I. Pulse-Position Modulation for Transmission Over Optical Fibers with Direct or Heterodyne Detection [J]. IEEE Transactions on Communications, 1983, 31(4):0–528.
7. Chen C C, Gardner C. Performance of PLL synchronized optical PPM communication systems [J]. Communications IEEE Transactions on, 1986, 34(10):988–994.
8. Georghiades C. Optimum Joint Slot and Symbol Synchronization for the Optical PPM Channel [J]. IEEE Transactions on Communications, 1987, 35(6):632–636.
9. Davidson F M, Sun X. Gaussian approximation versus nearly exact performance analysis of optical communication systems with PPM signaling and APD receivers [J]. Communications IEEE Transactions on, 1988, 36(11):1185–1192.
10. Davidson F M, Sun X. Slot clock recovery in optical PPM communication systems with avalanche photodiode photodetectors [J]. IEEE Transactions on Communications, 1989, 37(11):1164–1172.
11. Zwillinger D. Differential PPM has a Higher Throughput than PPM for the Band-Limited and Average-Power-Limited Optical Channel [J]. IEEE Transactions on Information Theory, 1988, 34(5):1269–1273.
12. Sugiyama H, Nosu K. MPPM: a method for improving the band-utilization efficiency in optical PPM [J]. IEEE Journal of Lightwave Technology, 1989, 7(3):465–472.
13. Barry J R. Sequence detection and equalization for pulse-position modulation [C]// IEEE International Conference on Communications. IEEE, 1994.
14. Malik D. Audeh, Kahn J M, Barry J R. Performance of pulse-position modulation on measured non-directed indoor infrared channels [J]. IEEE Transactions on Communications, 1996, 44(6): 654–6510.
15. David C, M. Lee, Kahn J M. Trellis-coded pulse-position modulation for indoor wireless infrared communications [J]. IEEE Transactions on Communications, 1997, 45(9):1080–1088.
16. Park H, Barry J R. Trellis-coded multiple-pulse position modulation for wireless infrared communications [C]// IEEE Global Telecommunications Conference. IEEE, 1999.
17. Shiu D S, Kahn J M. Differential pulse-position modulation for power-efficient optical communication [J]. IEEE Transactions on Communications, 1999, 47(8):1201–1211.
18. Nikolai F. Rulkov, Sushchik M M, Tsimring L S, et al. Digital communication using chaotic-pulse-position modulation [J]. IEEE Transactions on Circuits & Systems I Regular Papers, 2001, 48(12):1436–1444.
19. Aldibbiat N M, Ghassemlooy Z, Mclaughlin R. Dual header pulse interval modulation for dispersive indoor optical wireless communication systems [J]. IEE Proceedings - Circuits, Devices and Systems, 2002, 149(3):187–192.
20. Moision B, Hamkins J. Coded Modulation for the Deep-Space Optical Channel: Serially Concatenated Pulse-Position Modulation [J]. developmental biology,2005, 4(5)1–25.
21. Kiasaleh K. Performance of APD-based, PPM free-space optical communication systems in atmospheric turbulence [J]. IEEE Trans Commun, 2005, 53(9):1455–1461.
22. Wilson S G, Brandt-Pearce M, Cao Q, et al. Free-space optical MIMO transmission with Q-ary PPM [J]. IEEE Transactions on Communications, 2005, 53(8):1402–1413.
23. Ivan B. Djordjevic, Vasic B, Neifeld M A. Multilevel coding in free-space optical MIMO transmission with Q-ary PPM over the atmospheric turbulence channel [J]. IEEE Photonics Technology Letters, 2006, 18(14):1491–1493.
24. Maged F B, Moision B, Fitz M. Iterative Coded Pulse-Position-Modulation for Deep-Space Optical Communications [C]// Information Theory Workshop, 2008. ITW '08. IEEE. IEEE, 2008.
25. Sethakaset U, Gulliver T A. Performance of differential pulse-position modulation (DPPM) with concatenated coding over optical wireless communications [J]. IET Communications, 2008, 2(1):45–52.
26. Fang X, Khalighi M A, Bourennane S. Coded PPM and Multipulse PPM and Iterative Detection for Free-Space Optical Links [J]. IEEE/OSA Journal of Optical Communications & Networking, 2009, 1(5):404–418.
27. Wasiu O. P, Poves E, Haas H. Spatial Pulse Position Modulation for Optical Communications [J]. Journal of Lightwave Technology, 2012, 30(18):2948–2954.

28. Correia D G, Sales J C, Pinto P, et al. Analysis of the Performance of a PAM/PPM/OOK System Operating with OCDMA, under Nonlinear Optical Effects in Optical Fiber Propagation [J]. Journal of Optical Communications, 2016, 37(2):233–246.
29. Zohaib A. F, Mostafa H. A, et al. Practical implementation of duobinary pulse position modulation using FPGA and visible light communication [C]// 2017 IEEE 15th Student Conference on Research and Development (SCOReD). IEEE, 2018.
30. Sharma K, Grewal S K. Performance assessment of hybrid PPM-BPSK-SIM based FSO communication system using time and wavelength diversity under Atmospheric Turbulence [J]. Optical and Quantum Electronics, 2020, 52(10):430–455.
31. Zaiton A M, Eng C H, Jasman F. Pulse Position Modulation characterization for indoor visible light communication system [J]. Journal of Physics Conference Series, 2020, 5(1):1502–1512.
32. Yin B L, Wang F C. Relational Digital Phase Lock Loop slot synchronizer for PPM Laser Communication System [J]. OPTICAL COMMUNICATION TECHNOLOGY, 1998, 22(3): 211–218.
33. Su Y, Yao W C, Huang B X, et al. Design and Implementation of PPM Modulator Using EPLD [J]. OPTICAL COMMUNICATION TECHNOLOGY, 1999, 23(4):272–278.
34. Zhang M, Qiu K, Qiu Q, et al. Word Synchronization of Pulse Position Modulation [J]. Journal of University of Electronic Science and Technology of China, 2000, 29(4):381–383.
35. Du Y, Zou C Y. The Circuit Design for PPM Encoder and Decoder [J]. Journal of Guilin Institute of Electronic Technology, 2000, 20(3):5–11.
36. Wang J, Zhang Z X. PPM in Wireless Infrared Communication [J]. Telecommunication Engineering, 2000, 40(5):81–84.
37. Pang Z Y, Pu D Z, Zou C Y. Performance Comparisons of Several Modulation Schemes for Optical Wireless Communication [J]. Journal of Guilin Institute of Electronic Technology, 2002, 22(5):1–4.
38. Hu Z M, Tang J X. Digital pulse interval modulation for atmospheric optical wireless communications [J]. Journal of China Institute of Communications, 2005, 26(3):75–710.
39. Zhu L J, Zhang J X. Design of PPM Modulation System for Optical Wireless Communication Based on FPGA [J]. Optical Communication Technology, 2006, 30(12):57–510.
40. Cheng G, Wang H X, Long-Gang W U, et al. Performance Analysis of Modulation Scheme for Atmospheric Optical Wireless Communications [J]. Journal of China Academy of Electronics and Information Technology, 2007, 2(5):485–4810.
41. Zheng D, Sha X J, Yong W, et al. LDPC-coded optical PPM communication system over the atmosphere turbulence channels [C]. IEEE International Conference on Information & Automation, 2011.
42. Cheng G, Wang H X, Sun X M. A New Modulation Scheme of Optical Wireless Communications [J]. Chinese Journal of Lasers, 2008, 35(12):1914–1919.
43. Cheng G, Wang H X, Zhang T Y. Dual Pulse-Pulse Interval Modulation for Optical Wireless Communications [J]. Chinese Journal of Lasers, 2010, 37(7):1750–1755.
44. Xiang J S, Wu T, Huang S, et al. Slot Synchronization Aided by Serial Concatenated Pulse Position Modulation Code System [J]. Acta Optica Sinica, 2016, 36(8):63–70.
45. Xiang J S, Chen X L, Zhang P, et al. Clock Synchronization Technology for Pulse Position Modulation with Guard Time at Sampling Frequency of 1 Slot [J]. Chinese Journal of Lasers, 2018, 45(10):232–240.
46. Jiang T, Zhao L, Liu H Z, et al. Performance Improvement for Mixed RF–FSO Communication System by Adopting Hybrid Subcarrier Intensity Modulation [J]. Applied Sciences, 2019, 9(18):3724–3738.
47. Xu Y H, Guo Y, Qin L. Performance analysis of visible light communication system based on DPIM modulation [J]. China New Communications,2020, 22(17):27–29.
48. Ding D Q. Design of PPM for Laser Communication in Atmosphere [D]. Xi'an: Xi'an University of technology, 2005.
49. Qin L. Design of MPPM in Atmospheric Laser Communication [D]. Xi'an: Xi'an University of technology, 2008.

50. Ke X Z, Zhao L, Ding D Q. The Achievement of Time-slot and Frame Synchronization in Atmosphere Laser Communication [J]. Semiconductor Optoelectronics, 2007, 25(5):721–724.
51. Qin L, Ke X Z. Analysis of optical MPPM channel capacity without background noise [J]. Opto-Electronic Engineering, 2007, 34(7):107–111.
52. Huang L, Ke X Z, Wu P F. Design and simulation of OPPM-oriented TCM system [J]. Study on Optical Communications, 2008, 5(7):13–16..
53. Ke X Z, Chen J N. Performance comparison of various pulse position modulation in wireless laser communication [J]. Laser Technology, 2012, 36(1):67–76.
54. Ma L N. PPM Polarization Modulation in Atmosphere Laser Communication and Experiment Research [D]. Xi'an: Xi'an University of technology, 2011.
55. Zhao L. Research of DPPM for Laser Communication in Atmosphere [D]. Xi'an: Xi'an University of technology, 2006:13–29.
56. Huang L. The Application of TCM in Atmosphere Laser Communication System [D]. Xi'an: Xi'an University of technology, 2008.
57. Sun C M. Simulation and Research of DAPPM System for Laser Communication in Atmosphere [D]. Xi'an: Xi'an University of technology, 2011: 9–20.
58. Jia C. Research on DHPIM Modulation and Demodulation for Atmospheric Laser Communication [D]. Xi'an: Xi'an University of technology, 2011: 14–30.
59. Liu M P. Performances Analysis of PPMS' in Wireless Optica Communication [D]. Xi'an: Xi'an University of technology, 2013: 13–33.
60. Shi B Y. Experimental of 64-QAM Modulation in Optical Wireless Communication System [D]. Xi'an: Xi'an University of technology.
61. Tian X C. Adaptive Signal Processing to Receive the Strong Turbulence Atmosphere Laser Communication [D]. Xi'an: Xi'an University of technology, 2018.
62. Pei G Q. Performance Research on Multi-beam Transmission and Multi-Aperture Reception in Turbulent Atmosphere [D]. Xi'an: Xi'an University of technology, 2013.
63. Liu Z H. Research on Response Characteristics of Semiconductor Lasers and Photodetectors [D]. Xi'an: Xi'an University of Technology, 20110.
64. Zhang Y, Gao Y, Ke X Z. Analysis of spatial correlation of precoding indoor MIMO visible light communication system [J]. Opto-Electronic Engineering, 2020, 47(3):18–24.
65. Lin P, Wang T S, Ma W Z, et al. Propagation characteristics of 2.07 μm fiber laser in weak turbulence condition [J]. Opto-Electronic Engineering, 2020, 47(3):72–79.
66. Wang K, Xu Z Y, Li X S, et al. Analysis of space diversity method in modulating retro-reflector optical communication [J]. Opto-Electronic Engineering, 2020, 47(3):58–63.
67. Mao Y C, Wang H Q, Zhang Y, et al. Research status and development of optical spatial modulation technology [J]. Opto-Electronic Engineering, 2020, 47(3):37–48.
68. Ke X Z, Wu J L. Principle and application of wireless optical coherent communication [M]. Beijing: Science Press, 20010.
69. Ke X Z, Yang S J, Wu J L, Zhong X R. Research progress of adaptive optics in Optical Wireless Communication system for Xi'an University of Technology [J]. High Power Laser and Particle Beams, 2021, 33(08):30–52.
70. Ke X Z, Wu J L, Yang S J. Research progress and prospect of atmospheric turbulence for Optical Wireless Communication [J]. Chinese Journal of Radio Science, 2021, 36(03): 323–3310.
71. Ke X Z. Coding theory in wireless laser communication system [M]. Beijing: Science Press, 2011.
72. Ke X Z, Ding D Q. PPM synchronizer for atmospheric laser communication[P]. Xi'an: CN201066849 Y, 2011.
73. Ke X Z, Liu J. PPM polarization modulation and demodulation method for wireless laser communication[P]. Xi'an: CN101110649A,2009.
74. Ke X Z, Liang J Y, Xu D S, et al. Research progress of pulse position modulation technology in optical wireless communication [J]. Opto-Electron Eng, 2022, (03):3–21.

Communication Lasers and Their Modulation Technology

4

Contents

Introduction 86
Lasers and Optical Wireless Communication 86
Wireless Communication Systems 86
Laser Structure 87
Laser Modulation Modes 88
Laser Research Progress and Present Situation 88
Solid-State Lasers 88
Gas Lasers 90
Semiconductor Lasers 91
Fiber Lasers 96
Low-Coherence Lasers 97
Ultraviolet Lasers 97
Modulation Technology 98
Internal Modulation 101
External Modulation 103
Lasers and Modulation Technology 108
Laser-Characteristic Optimization 109
Modulation Experiment 119
Internal Modulation 119
Electro-optic Modulation 120
Summary 124
References 125

Abstract

In the field of optical communication, laser and its modulation bandwidth and modulation rate are very important for optical wireless communication. On the basis of reviewing the development history of communication laser and its modulation technology, this chapter introduces the work done by Xi'an University of Technology in this field; discusses the internal and external factors of nonlinear distortion of semiconductor laser; simulates the effects of various parameters of response characteristics, linearization compensation method, and

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_4

power control system on the output light source of light emitting device; and analyzes the influence of temperature fluctuation on the output power of laser and modulation device.

Keywords

Optical wireless communication · Laser · Direct modulation · External modulation · Response characteristic

Introduction

The laser is an important light-emitting device used in optical wireless communication. Owing to its good monochromaticity, directivity, coherence, and intense brightness, the laser is an ideal choice for a system-emission light source, and it is used as a carrier in optical communication.

Light generation cannot be separated from a suitable light source. Not all light sources can be used in optical communication. The light sources used in optical communication must satisfy a series of requirements, such as adequate wavelength, power, spectral linewidth, modulation bandwidth, integration, and cost. Modulation is the process of converting a signal. After modulation, the information from the source signal is carried in the optical signal, which is transformed into a form that is convenient for transmission through the channel.

Laser modulation refers to a change in one or several characteristic parameters of an optical signal (such as the laser intensity, frequency, phase, or polarization) through a specific method, and the change rule is specific and controllable. Laser modulation is an important application technology that has been widely used in laser communication, sensing, displays, and processing. High-power light sources and high bit-rate modulation technologies are continuous development directions.

Lasers and Optical Wireless Communication

Wireless Communication Systems

An optical-communication system model is illustrated in Fig. 4.1. The system can be divided into three blocks: an optical transmitter, a channel, and an optical receiver. The optical transmitter includes coding equipment, a modulator, a laser, an optical signal amplifier, and a transmitting antenna. Transmission channels can be divided into atmospheric, indoor, underwater, and ultraviolet light-scattering channels. The optical receiver includes receiving antennas, detectors, demodulators, and decoders.

In optical wireless communication, the transmitter first modulates the information generated by the source onto an optical carrier, and then the carrier is transmitted through the atmosphere or free space. An acquisition, pointing, and tracking (APT) system controls the laser beam and ensures a smooth communication link.

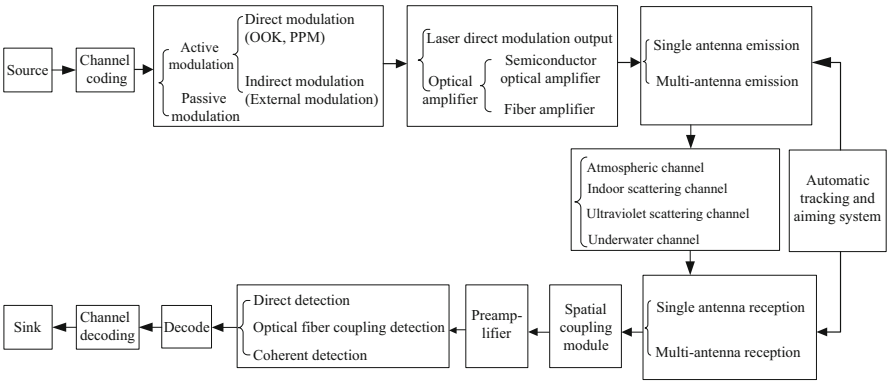


Fig. 4.1 Optical wireless communication model [1]

Subsequently, the optical signal from the transmitter is received by the optical signal antenna, and the received signal light is coupled onto an optical fiber using a space optical-fiber coupling device. Occasionally, the signal light is amplified by a preamplifier before being converted to electricity. Then, the baseband signal is recovered by the demodulator and the original baseband signal is obtained by the decoder.

Laser Structure

A semiconductor laser uses semiconductor material as the gain medium. It emits light as photons transition between semiconductor energy bands and usually takes a natural cleavage plane as the resonant cavity. Therefore, it has the advantages of a wide wavelength coverage, small volume, stable structure, strong radiation resistance, various pumping modes, high yield, good reliability, and easy high-speed modulation. Unfortunately, it has a poor output-beam quality, large beam-divergence angle, asymmetric spots, poor spectral purity owing to the influence of interband radiation, and a difficult preparation process.

As shown in Fig. 4.2, a laser is generally composed of a working substance, excitation source, and resonant cavity. Working substances can be solids (including various semiconductors), liquids, or gases, with different types and pumping methods. The substance must able to invert the particle number. Total and partial reflectors are placed in parallel to form a resonant cavity, which reflects the photons generated by the active gain layer and oscillates coherently to amplify the light. External excitation supplies energy to the working substance, enabling particles at a low energy level to jump to a higher energy level. Therefore, a laser with high photon degeneracy must satisfy three conditions: particle-number inversion, appropriate resonant cavity, and a threshold condition.

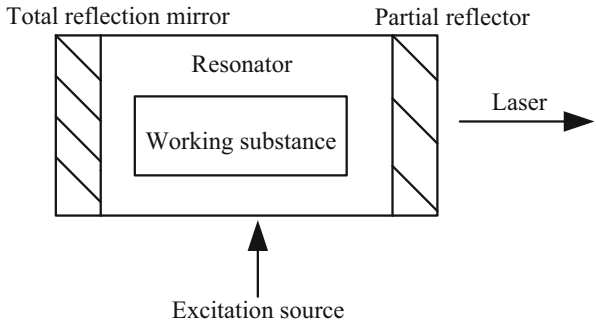


Fig. 4.2 Laser structure [2]

Laser Modulation Modes

Modulation implies that the information (voice, data, etc.) to be transmitted is transmitted through a certain medium by changing the carrier-signal parameters. In optical wireless communication, modulation means that information is loaded onto a laser or other carrier using an electro-optical conversion process, and the signal is converted into a form that is convenient for transmission in the channel.

Modulation can be divided into active and passive, depending on whether the light source and modulation signal are on the same end of the transmission. In active modulation, both the light source and modulated signal are at the transmitting end, whereas in passive modulation, the light source is at the transmitting end and the modulated signal is at the receiving end. Active modulation can be classified into direct and indirect types. The modulation of the laser power is direct, and the modulation of the light source emitted by the laser is indirect; it is also known as external modulation [1].

Laser Research Progress and Present Situation

Based on their different working media, lasers are subdivided into solid-state, gas, semiconductor, and fiber. Their development in communication is discussed.

Solid-State Lasers

In a solid-state laser (SSL), the working medium is mainly a solid metal material. In 1917, Einstein explicitly proposed stimulated radiation, which provided the main theoretical basis for the later discovery of lasers [3]. In 1960, Maiman successfully developed the first chromium (Cr)-doped ruby solid-state laser pumped by a pulsed xenon lamp [4].

Since the 1980s, solid-state lasers have shown clear directions, including all-solid-state, tunable, high-power, and new wavelength [5]. Research on all-

solid-state 532-nm green lasers with high power and miniaturization has played an important role in ocean exploration, underwater communication, and other fields.

With the development of optical-fiber communication networks, particularly after the wide application of dense wavelength-division multiplexing (DWDM) technology in the optical-communication field, researchers have proposed higher requirements for lasers as communication light sources. Currently, fixed-wavelength lasers are traditionally used as light sources. If multiple wavelengths are output, the cost and volume of the laser increase correspondingly; this inevitably limits the flexibility of optical network expansion.

If a wavelength-tunable laser is used, instead of a traditional fixed-wavelength laser, the high cost of the light-source configuration and maintenance in the DWDM system can be reduced, the wavelength resources can be fully utilized, the production cost and management complexity can be reduced, and the cost performance of the network system can be improved.

Next, we introduce the tunable laser whose tuning range is limited by the bandwidth of the gain medium. Tuning involves controlling the loss in the laser cavity to minimize the loss of a specific laser wavelength. Not all wavelengths for a given application can be generated from a laser using the required parameters (such as the function or energy). Wavelengths are typically converted when other laser wavelengths are required.

To create wavelength-tunable lasers, according to the demands of the late 1990s, the wavelength-tuning range should be approximately 40–60 nm. Let us consider nonlinear-wavelength conversion. In 2006, a side-pumped 1064-nm neodymium-doped yttrium–aluminum–garnet (Nd:YAG) rod V-shaped cavity laser was developed by Mukhopadhyay et al. in Pakistan. Potassium titanyl phosphate ($KTiOPO_4$, KTP) was used as a nonlinear frequency-doubling crystal, and a green light with a beam-quality factor of 20 was obtained by frequency doubling [6].

In 2009, a double-headed series structure based on side pumping was used by Xu et al. Lithium triborate (LBO) was used as a nonlinear frequency-doubling crystal. The output power of the continuous green light reached 40 W, and the electro-optical conversion efficiency was 5.0% [7]. In 2011, Singh et al. doubled the intracavity frequency of a diode-pumped Nd:YAG laser in a Z-shaped cavity, showing a continuous green power of 82.5 W, corresponding to a pump-to-green-light conversion efficiency of 11% [8].

In 2012, Ji et al. adopted a compact and unstable V-shaped cavity approximately 300 mm in length. Using the green laser generated by the second harmonic generation (SHG) in the cavity, the light–light conversion efficiency from the diode to green light was 110.7% and the infrared–green conversion efficiency was greater than 80% (81.8%) [9]. The development of all-solid-state lasers can compensate for the shrinking market of traditional solid-state lasers.

Solid-state lasers generally have short wavelengths (ultraviolet and deep ultraviolet), short pulse widths (picoseconds and nanoseconds), and high peak power. To produce an ultraviolet solid-state laser, the pump light source in the laser irradiates the enhancement medium to invert the particle beam; a fundamental infrared light is formed and oscillates in the resonant cavity. After one or more frequency-doubling

cycles in the nonlinear crystal cavity, the required ultraviolet laser is finally output from the resonant cavity through transmission and reflection.

Nd:YAG and neodymium-doped yttrium orthovanadate (Nd:YVO_4) are both common enhancement dielectric crystals. A commonly used method to enhance the resonant cavity is to pump an 808-nm-wavelength Nd:YVO_4 laser crystal into the cavity for frequency doubling to output a 532-nm-wavelength green light, and then send it to the enhancement resonant cavity for frequency doubling to output a 266-nm-wavelength deep ultraviolet laser. The input threshold of a fundamental green light can be as low as 215 mW.

The 28-nm wavelength band has a very important application in ultraviolet resonance Raman spectroscopy. In 2022, Zhao et al. adopted a V-shaped intracavity frequency-doubling and lens-focusing mode intracavity frequency-quadrupling structure. Based on an Nd:YVO_4 quasi-three-level laser system, a 914-nm fundamental-light cascaded nonlinear optical frequency was converted to obtain a continuous-wave (CW) 228-nm laser output [10]. In general, compared to the pulse-operation mode, a continuous-operation laser requires a higher average power and better beam quality to realize the extra-cavity frequency doubling of the lens focus.

Solid-state lasers generally have short wavelengths (ultraviolet and deep ultraviolet), short pulse widths (picosecond and nanosecond), and high peak power. However, their beam quality is affected by temperature; therefore, thermal management is very important for the overall efficiency of the laser system. Commonly used cooling methods include spray (water) cooling [11], solid (diamond) cooling [12], direct heat dissipation of crystal cross sections [13], forced-convection cooling, semiconductor refrigeration, and heat-pipe heat-conduction technology [14].

Further development depends on improving the thermal management, as well as finding new nonlinear materials and obtaining the best wavelength required for specific applications [15].

Gas Lasers

The earliest gas lasers, which use gas as a gain material, were He–Ne gas lasers [16]. In 1965, Patel et al. successfully discharged a N_2–CO_2–He gas mixture, which was the first electrically excited CO_2 laser, and produced a laser with high continuous output power [17].

The acousto-optic Q-switched Q CO_2 laser satisfies the application requirements for space communication. In 1991, Yang et al. developed an acousto-optic Q-switched continuous direct-current (DC)-discharge-excited waveguide CO_2 laser. The acousto-optic medium was a germanium single crystal, and the pulse-repetition frequency range was 0–200 kHz [18].

Since 2009, Professor Xie Yijiang et al. have studied the specific technical parameters of acousto-optic Q-switching, using the rate equation of a Q-switched pulse laser, and manufactured an acousto-optic Q-switched Q CO_2 laser with tunable coding output [18–20]. In 2012, based on their achievements, Xie et al. adopted the

method of inserting an acousto-optic modulator into the cavity. The adjustable laser-pulse frequency range was 1 Hz–100 kHz. When the pulse frequency was 1 kHz, the pulse width was 156 ns, and the peak pulse power was 10 kW [21].

In 2019, Pan et al. successfully used the external cavity method to study modulator characteristics and successfully integrated the optical-path fast-deflection technology of the double acousto-optic modulator with a self-collimating grating method. This enabled the fast tuning of the CO_2 laser over the entire band and the output of the same optical path, while reducing the CO_2 laser and wavelength switching time to milliseconds [22].

Gas lasers were the first lasers to be applied, and they still occupy an irreplaceable position in lithography and other fields. However, owing to their large size and difficult maintenance, the market share of gas lasers in China has been less than 5% in recent years, and they are no longer mainstream lasers.

Semiconductor Lasers

Semiconductor (laser diode, LD) lasers, whose working medium is semiconductor materials, are usually doped with III–V compounds, such as gallium arsenide (GaAs), indium phosphide (InP), gallium nitride (GaN), etc. The lasers used for optical communication are primarily semiconductor lasers. Currently, the development can be roughly divided into two categories: information and power.

Information lasers are mainly used in the field of information technology, such as distributed feedback and dynamic single-mode lasers (DFB-LD) for optical-fiber communication and optical switching systems, tunable lasers with a narrow linewidth, and visible-light wavelength lasers (405 nm, 532 nm, 635 nm, 650 nm, and 670 nm) for information-processing fields, such as optical disks. They exhibit a single frequency, narrow linewidth, and high speed.

Power lasers are used as a pumping source in applications such as pumping solid-state lasers. This has led to a breakthrough in high-power semiconductor lasers, which was marked by a significant increase in their output power. The basic structure of a semiconductor laser is shown in Fig. 4.3.

According to the relationship between the beam exit direction and the substrate surface, semiconductor lasers can be divided into edge-emitting lasers (EEL) and vertical-cavity surface-emitting lasers (VCSEL). Figure 4.4 shows their structures and characteristics.

Edge-Emitting Semiconductor Lasers

Edge-emitting lasers include distributed-feedback lasers (DFB), Fabry–Perot (F–P), distributed Bragg reflectors (DBR), and so forth. They can be used as direct-modulation lasers or coupled with optical fibers or crystals to manufacture diode-pumped solid-state lasers. They have a wide wavelength range and high photoelectric efficiency, and their applications include optical communication, optical storage, material processing, medical treatment, sensing, etc.

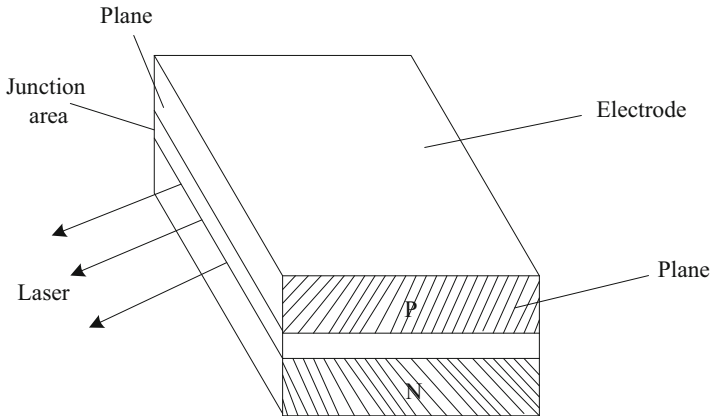


Fig. 4.3 Basic structure of a semiconductor laser [1]

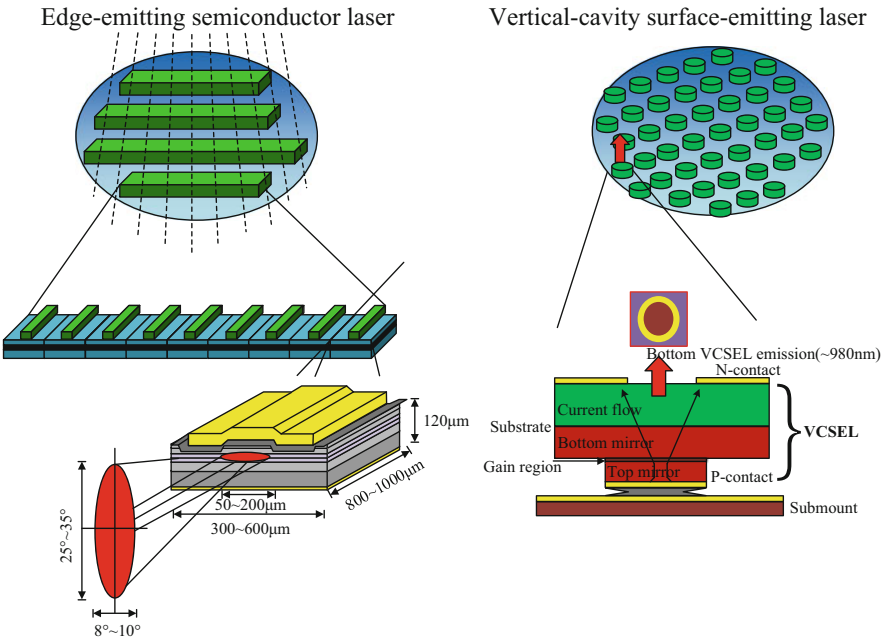


Fig. 4.4 Basic structure and characteristics of edge-emitting and surface-emitting lasers [23]

With the increasing demand for capacity, speed, and stability in optical-fiber communication, semiconductor lasers must have a lower threshold, better dynamic single-longitudinal mode, higher modulation rate, and higher reliability. The difference between F–P, DFB, and DBR lies in the longitudinal mode, which is determined by the design of the resonant cavity. The earliest F–P cavity semiconductor

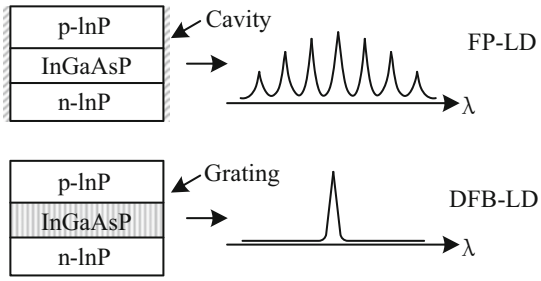


Fig. 4.5 Structural-comparison diagram of FP-LD and DFB-LD

lasers could not meet the demands of a single-mode optical-fiber communication system because of their multiple longitudinal modes.

To solve this problem, Bragg-grating technology was introduced into LD. When the Bragg-grating laser is used instead of the F–P cavity, only half of the grating is needed to adjust the phase by current, that is, the output wavelength can be tuned by the magnitude of the current. Alternatively, a Bragg-grating structure can be applied to the active region, which is the widely used DFB laser. These two types of lasers can stabilize narrow emission linewidths.

For DFB lasers, only one grating can be used for wavelength selection, and they are widely used in medium- and long-distance transmissions, including transmission networks, wireless base stations, and data-center internal interconnections. They are some of the most important light sources used in communication systems. Figure 4.5 shows a structural comparison of the FP-LD and DFB-LD.

From 1971 to 1972, Kogelnik of Bell Laboratories proposed the basic theory of the DFB laser [24, 25]. In 1973, a semiconductor DFB laser was developed [26]. At the end of 1988, a DFB laser with a ridge waveguide at the 1.55-μm band was developed in China, and the side-mode suppression ratio exceeded 30 dB. Direct-modulation DFB-LDs are frequently used in high-speed short-distance optical-fiber communication technology; their research progress will be discussed in the internal-modulation section.

Compared to current quantum-well devices, the future development direction of DFB-LDs will be a combination of large-scale functional integration and various material systems. For example, quantum-dot gain media can significantly improve the manufacturing efficiency, producing DFB lasers with smaller volumes and better performance to meet future market needs [27]. Mid-infrared continuous-wave (CW) semiconductor lasers are very important light sources, especially in spectroscopy. GaSb-based interband emitters can be used as effective seed sources for tunable semiconductor laser spectra, both in external-cavity construction and spectrum-selective distributed feedback, and have relatively low power.

Previously reported semiconductor lasers, based on multiple W-quantum-well active areas, could only output pulsed light, and the working mechanism was due to the hole-tunnel carrier-transport mode in multiple 2-nm-thick InAs layers, which led to the low utilization rate of multiple W-quantum-well active areas. Researchers at Stony Brook University proposed an improved design for W-quantum wells without

an InAs layer. They designed and constructed a CW semiconductor laser based on a three-layer GaInAsSb quantum well at room temperature.

The working wavelength of the laser is flexible and adjustable, and can be independently controlled by the effective electron and hole quantum-well widths. The improved design scheme adopts a quantum well with a 10–12-nm width, so that its working wavelength can reach more than 3.2 μm, and the CW working range at room temperature can be close to 3.3 μm or 3.4 m.

Laser-pumped praseodymium-doped solid-state lasers have also made progress. Researchers in the Department of Electronic Engineering of Xiamen University proposed a praseodymium-doped diode-pumped laser based on $LiYF_4$, which works at 695.8 nm. In the experiment, the high gain conversion was suppressed at wavelengths of 640, 670, 698, and 721 nm. They used InGaN as the laser diode in the pump source, and the output power of the σ polarized light at the central wavelength of 695.8 nm was 51 mW. Its slope efficiency was 13.6% and the threshold of the absorbing pump power was 229 mW. Researchers used a yttrium–lithium–fluoride (YLF) praseodymium-doped laser to obtain the output of the σ polarized light at a central wavelength of 695.8 nm.

Infrared, red, green, and blue lasers have been used practically. In recent years, the industry has generated a demand for more powerful ultraviolet lasers. In February 2020, Professor Susumu Iwatani, Faculty of Science and Technology, Meishiro University, Japan, announced the invention of the world's first mid-wavelength ultraviolet semiconductor laser, using independently developed semiconductor technology and new methods [28].

Compared with gas and solid-state lasers used in the same wavelength range, semiconductor lasers can provide miniaturization, low power consumption, long service life, and low cost. This achievement will open up new possibilities for optical science, which can not only be applied to industry but also be expected to solve problems in other fields, such as chemistry, environment, medical treatment, and biological science.

Surface-Emitting Semiconductor Lasers

In the general structure of a vertical-cavity surface-emitting laser (VCSEL), the upper and lower DBR structures are used as the emitting cavity and gain region in the middle, respectively. A dynamic single-mode operation is ensured, based on the DBR structure. The volume of the gain region is small, the threshold current is low, and the optical-fiber coupling efficiency is extremely high. The outgoing Gaussian beam is astigmatic and forms a high-density two-dimensional array, making it important for applications in optical interconnection and optical information processing [29].

The VCSEL for the optical module includes a multi-transverse mode and a single longitudinal mode. For medium- and low-speed data-communication networks, 850-nm VCSEL is the mainstream short-distance data-communication method because of its low cost, short active optical cavity, and high modulation rate. In 1977, the concept of a VCSEL was first proposed by Iga in Japan [30], with a 900-mA threshold current. In 1988, the VCSEL device successfully produced

continuous lasing at room temperature in the 850-nm band [31], which was a key step towards practical applications. In 1996, Lear et al. used an oxide-limited VCSEL for optical communication. It limited the light field and reduced the threshold current. The main parameter required was the highest-possible modulation speed.

In 2008, Finisar Company successfully fabricated a shortwave communication VCSEL in the non-return-to-zero (NRZ) modulation format using unstrained GaAs quantum wells in the active area. At room temperature, the rate was 30 Gbit/s and the bandwidth reached 19 GHz [32]. In 2015, IBM Corporation cooperated with Chalmers University of Technology in Sweden to integrate driver circuits, decoupling capacitors, and VCSEL, and introduced equalization technology. The error-free data-transmission rate was 71 Gbit/s at 25 °C and 50 Gbit/s at 90 °C [33].

From 2001 to 2015, the modulation rate of VCSEL NRZ increased from 20 to 71 Gbit/s [32–37]. After 2015, the modulation rate and bandwidth increased. Combined with multiwavelength or complex modulation, it was applied at 100, 200, and 400 GHz.

Currently, the transmission rate of a shortwave GaAs-based 850-nm VCSEL exceeds 100 Gbit/s. PAM4 is a type of pulse-amplitude modulation (PAM) technology. PAM is a popular signal-transmission technology after NRZ and is representative of multilevel modulation technology. PAM has been widely used in the field of high-speed signal interconnections. Under the same baud-rate conditions, the bit rate of the PAM4 signal is twice that of the NRZ signal, the transmission efficiency is doubled, and the cost can be effectively reduced.

In 2020, Lavrencik et al. [38] used forward error correction and a receiver equalizer in their system to improve the data rate. Experiments showed that the data rate of four-level pulse-amplitude modulation (PAM4) exceeded 160 Gbit/s. In 2021, Zuo [39] combined a nonlinear equalizer and a noise-cancellation algorithm in digital signal processing to alleviate the intersymbol interference and nonlinear factors caused by PAM4. Consequently, the transmission rate of a VCSEL on 100-m optical mode (OM4) multimode fiber reached 200 Gbit/s. Table 4.1 lists the chip research progress of the 850-nm VCSEL for data communication.

Because the information-transmission distance of shortwave VCSEL optical communication is limited to several hundred meters by modal dispersion, it has

Table 4.1 Research progress of the 850-nm VCSEL chip for data communication [40]

Research institution	Modulation mode	Rate (Gbit/s)	Year	Reference
IBM	NRZ	20	2001	[34]
Finisar	NRZ	30	2008	[32]
Chalmers U.	NRZ	40	2010	[35]
Chalmers U.	NRZ	57	2013	[36]
IBM	NRZ	64	2014	[37]
IBM–Chalmers	NRZ	71	2015	[33]
Lavrencik	PAM4	160	2020	[38]
Zuo	PAM4	200	2021	[39]

been necessary to develop a long-wavelength band with less loss for investigating high-speed VCSEL, such as 1310-nm and 1550-nm band systems [41–44]. It is possible to extend the communication distance to over 20 km.

In contrast to the perfected preparation process of 850-nm VCSEL, at present, the research on long wavelengths mostly lies in optimizing the device structure, which mainly involves preparing a DBR with a high refractive index and high thermal conductivity, forming an effective current-limiting structure, and solving thermal problems [45].

Fiber Lasers

In a fiber laser, the gain substance is called the gain fiber, and its output wavelength is mainly determined by doped rare-earth elements. For example, a Yb^{3+}-doped glass fiber outputs a 1064-nm laser and an Er^{3+}-doped glass fiber outputs a 1310-nm or 1550-nm laser. The fiber modules are divided into single mode and multimode. The fiber core of a single-mode laser is thinner and the beam quality is better than that of a multimode. The energy distribution is Gaussian, and the energy density in the middle is the highest. A three-dimensional figure shows a sharp, round mountain peak.

The fiber core of a multimode laser is thicker, the beam quality is worse, and the energy distribution is more normal than that of single-mode. The three-dimensional image is an inverted cup, and the edge of a multimode laser is much steeper than that of a single-mode. Owing to their excellent beam quality and flexible power-expansion capability (fiber combiner), fiber lasers have developed rapidly and can realize CW operation and high-repetition-rate pulse operations [46].

In 1961, Snitzer [47] of the United States proposed the concept of a fiber laser. In 1966, Gao theoretically analyzed the possibility of using optical fibers as a new type of communication transmission medium and proposed using impurity-free and transparent glass fibers to transmit laser signals. If the loss is as low as 20 dB/km, long-distance optical communication can be realized, promoting the development of optical fibers in the communications field [46].

Mode-locked fiber lasers with tunable wavelengths have important applications in spectroscopy, optical communications, and sensing. This research topic has become popular in recent years. At present, the most common and effective method to build a wavelength-tunable laser uses a filter with wavelength selectivity and a mode-locked device with a nonlinear saturable absorption effect. In 1989, Hanna et al. demonstrated the first mode-locked erbium-doped fiber laser using a bulk lithium–niobate (LN or $LiNbO_3$) phase modulator with a 4.5-W output power and 37-ps mode-locked pulses [48].

In 2015, a mode-locked erbium-doped fiber laser developed by Li et al. adopted nonlinear polarization rotation technology and produced a pulse with a 34.3-fs width [49]. In 2022, Mao et al. proved that birefringence could also be used to modulate the spectrum and time-domain shape of pulses in the positive-dispersion region. This type of fiber laser can generate harmonic mode-locked birefringence management

solitons with an adjustable pulse-energy range exceeding 15 times. It provides an effective method to create a pulse source with a high repetition rate and low power consumption [50].

Since the birth of lasers, the early stage mainly involved improving the basic theory. In the last 30 years, lasers have developed rapidly under the huge driving effect of practical applications. Currently, the research hotspots of communication lasers are high power, short pulses, and wavelength extensions to the mid-infrared region. Research teams worldwide are constantly improving laser performance and expanding their fields of application.

Driven by the application of pumped solid-state and fiber lasers, the output power of semiconductor lasers has increased significantly, and semiconductor optoelectronic devices are developing in the directions of longer and shorter wavelengths, higher power, and higher working frequency; however, further performance improvement is still a goal [40].

Low-Coherence Lasers

Lasers are widely used because of their high coherence, directivity, and monochromaticity. However, in the laser-display and biological-imaging fields, the speckle phenomenon caused by high laser coherence seriously affects the imaging quality. Goodman concluded [51] that shaking screens and laser arrays could reduce speckle; however, these methods have their own problems, such as mechanical instability, huge volume, and high cost. The most direct and thorough way to eliminate speckles is to use a low-coherence laser.

Through innovations in the geometry of the laser cavity and the laser-oscillation feedback mechanism, low-coherence lasers, developed by regulating the number of laser modes and the spatial distribution of the mode field, are receiving increased attention. Over the past 30 years, low-coherence lasers have developed rapidly, and random lasers, degenerate-cavity lasers, chaotic-cavity semiconductor lasers, near-concentric-cavity semiconductor lasers, and dumbbell-shaped-cavity semiconductor lasers have successively emerged. However, low-coherence lasers are still in the laboratory-research stage, and many urgent problems remain to be solved, including poor directivity, low power, continuous lasing, and few means of adjusting the coherence. Therefore, the application of low-coherence lasers requires further investigation.

Ultraviolet Lasers

Ultraviolet lasers are mainly classified into gas and solid-state ultraviolet lasers. Under the action of the pump source, the working medium reaches the excited state by absorbing external energy, and the gain is greater than the loss through the particle-number inversion. Hence, the light is amplified, and part of the amplified-

light feedback continues to be excited, thus creating an oscillation in the resonant cavity to generate laser light.

The gas medium mainly uses a pulse or electron-beam discharge, and gas particles are excited from low to high energy levels by mutual collisions between electrons, resulting in a stimulated transition and thus, producing an ultraviolet laser. A solid medium is a type of ultraviolet laser that radiates outward after one or more frequency conversions in the form of nonlinear frequency-doubling crystals.

Quasimolecular and all-solid-state ultraviolet lasers are commonly used for laser processing. Ultraviolet gas lasers mainly include excimer lasers, argon–ion lasers, nitrogen-molecular lasers, fluorine-molecular lasers, and He–Cd lasers. The advantages of all-solid-state ultraviolet lasers are their convenience, small size, high reliability, and stable operation. See Table 4.2 for the relevant parameters of ultraviolet lasers. Table 4.3 presents a comparison of the properties of common light sources used in optical communications.

Modulation Technology

Light-modulation modes include amplitude, frequency, phase, and pulse modulation. Depending on the type of interaction between light and matter, it can be shown as the control effect of electricity, sound, or magnetism on light and can also be divided into analog and digital modulation. Choosing an efficient and reliable modulation technology guarantees a high transmission rate and low bit-error rate. This section focuses on internal and external modulation in active modulation. As

Table 4.2 Parameters of ultraviolet lasers

Ultraviolet laser classification	Spot shape and wavelength	Characteristics
Solid-state Q Nd:YAG laser	Wavelength: 353 nm Spot: Round (the energy decreases gradually from the center to the edge) Focus magnitude is 10 μm	High temperature sensitivity, which can only be stabilized after a cold-start period. High-repetition frequency, suitable for small-size processing
Excimer laser (gas laser)	Wavelength: Depends on the gas type Spot: Rectangular (Mask technology can produce different shapes of spots)	The processing details can be as small as a few microns, and the distance between the focusing mirror and the workpiece can be 50–100 mm
Metal-vapor laser (mainly copper laser vapor)	Wavelength: Copper vapor with 511-nm and 578-nm wavelengths generates ultraviolet light with 255-nm, 271-nm, and 281-nm wavelengths through frequency mixing and doubling. The light beam has a Gaussian distribution	The application range is the same as that of a solid-state ultraviolet laser

Table 4.3 Comparison of the properties of common optical-communication light sources

Type of light source	Typical wavelength (nm)	Line width	Output power	Beam shape	Fiber coupling	Transmission distance	Modulation bandwidth	Temperature stability	Cost
LED	850, 1300	>20 nm	<0 dBm	Large divergence angle	Simple	Short (<100 m)	<500 MHz	<0.5 nm/°C	Extremely low
VCSEL Laser	850	1–2 nm	<0 dBm	Circular spot, vertical emission	Simple	Short (<500 m)	>10 GHz	<0.1 nm/°C	Low
Fabry–Perot Laser	1280–1330 1480–1650	2–10 nm	>0 dBm	Elliptical spot, side emission	Complex	Medium (<10 km)	>10 GHz	<0.1 nm/°C	Medium
DFB Laser	1280–1330 1480–1650	<0.2 nm	>0 dBm	Elliptical spot, side emission	Complex	Long (>100 m)	>20 GHz	<0.1 nm/°C	High

Table 4.4 Differences between internal and external modulation [52]

Modulation mode		Modulation bandwidth	Chirp factor	Drives current or voltage	Impedance	Conversion relationship
Internal	DFB MQU-DFB	About 12 GHz <20 GHz	3–5 About 2	>50 mA >50 mA	4–8 Ω 4–8 Ω	Linear Linear
External	EA M-Z	About 40 GHz About 50 GHz	<1 0	About 1.5 V About 5–10 V	About 0.1 pF R = 50 Ω	exp(v) $\cos^2(v)$

listed in Table 4.4, external modulation is more suitable for high-speed transmission systems (>10 Gbit/s).

Currently, DFB lasers are commonly used as single-mode semiconductor light sources for optical communication. The modulation bandwidth and spectral purity are satisfactory, and they can be used for high-speed and long-distance transmissions. A general DFB laser modulates the light intensity by changing its driving current. A laser adopting this modulation mode is called a direct modulated laser (DML).

The DML modulation method is simple and low cost; however, it has several disadvantages for high-performance applications:

1. The modulation bandwidth is limited by its own relaxation-oscillation frequency, which restricts the highest data rate it can support.
2. A change in current will change the internal state of the laser, thus causing a wavelength shift.
3. To balance the modulation rate and performance, the laser is not completely off or on during modulation.
4. Its extinction ratio is limited, which limits the modulation rate, modulation quality, and transmission distance of the laser.

To further improve the modulation rate and quality, the optical signal can be modulated by an external modulator, that is, an electroabsorption modulated laser (EML). The EML is an external electroabsorption modulator (EAM) added to a common DFB light source. Mach–Zehnder modulators (MZM) are used for higher-performance applications.

Consider an EAM modulator as an example. The laser light source is always on, and its working current and output power are stable. The electrical signal modulates the power of the optical signal by controlling the voltage of the EAM and changing the light-absorption ratio, thereby avoiding the influence of relaxation oscillation and wavelength drift. If the modulation bandwidth, linearity, and extinction ratio of EAM devices can be improved, the modulation quality can also be improved. The modulation bandwidth of an EML laser can exceed 30 GHz and its wavelength stability is very high.

An MZM may also be used for applications with higher data rates (e.g., >50 Gbaud) or longer transmission distances (>40 km). The MZ modulator is typically composed of lithium niobate ($LiNbO_3$) and includes a pair of coplanar phase modulators on one plane. The input optical signal is divided into two branches, on average, and then combined using two phase modulators.

Lithium niobate has an electro-optic effect; that is, when an electric field is applied to its crystal, its refractive index changes. Therefore, when an electrical modulation signal is applied to one branch of the MZ modulator, the phase of the light passing through this branch will change, owing to the change in refractive index. However, when the light of one branch with a phase change is combined with another branch, the combined light signal may be strengthened or cancelled, depending on the phase difference between the two paths of light, thus modulating the light intensity.

Internal Modulation

Internal modulation, also called direct modulation, changes the laser-oscillation parameters by using the laws of modulation signals to change the characteristics of the laser output. Figure 4.6 shows two methods of realizing internal modulation. The first is to directly adjust the laser pump source (for example, a solid-state laser, which can be changed by the power, frequency, and duty ratio of the pump output); this directly affects the final output of the laser and changes the output laser light intensity, as shown in Fig. 4.6a.

The second is to place the modulation element in the resonant cavity. The output characteristics of the laser can be changed by changing the parameters in the cavity, as shown in Fig. 4.6b. Q-switching is a method for realizing internal laser modulation.

As a direct-current injection device, the change in the driving current of a semiconductor laser has an important impact on the output quality and output-power stability. Figure 4.7 shows the bias-T structure of a direct-modulation laser, which can modulate the data to the linear range of the laser. The bias and modulation currents are simultaneously applied to the laser. Direct-modulated lasers (DML) mainly include the Fabry–Perot laser (FP-LD) and distributed-feedback laser (DFB-LD). Owing to its low-intensity noise and high linearity, the DFB-LD has

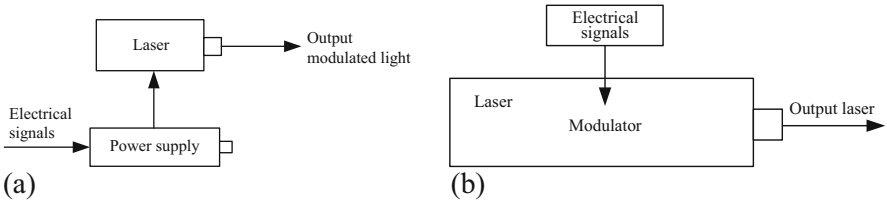


Fig. 4.6 Internal modulation systems [53]

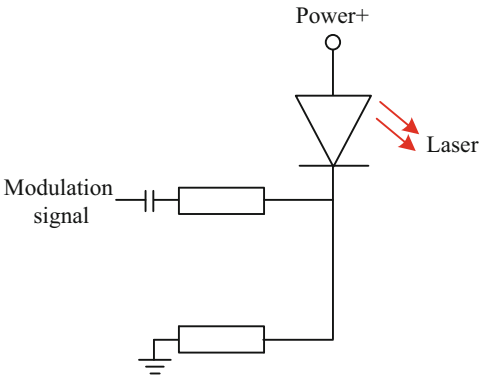


Fig. 4.7 Bias-T structure of a direct-modulation laser

become a key light source in communication systems. Further improvement in the modulation bandwidth of DFB-LD has always been a direction for researchers.

In 1964, Konnerth et al. were the first to study the direct modulation of semiconductor lasers [54]. In the same year, Goldstein et al. modulated GaAs laser microwaves with a modulation bandwidth of 2 GHz [55].

Since the 1990s, the optimization of high-speed direct-modulation lasers has focused on the material system and structural design of the active-gain region. Using a strained (compensated) multi-quantum-well structure in the active region and p-type doping in the active region can yield a very-high differential gain and more uniformly distribute the carriers in the quantum well.

In 1992, Bell Laboratories adopted an active-gain layer with multiple quantum wells and introduced strain into the structure, which effectively improved the differential gain; the modulation bandwidth reached 25 GHz [56].

In 1997, OKI Company realized that the number of quantum wells would affect the output power of lasers and designed 20 pairs of quantum-well DFB lasers with a 30-GHz bandwidth [57].

The selection criteria for the materials were good temperature characteristics and a high differential gain. In 2019, Sasada et al. used AlGaInAs quantum-well materials in the active region to obtain a 1.3-μm directly modulated laser with a 34.2-GHz bandwidth at 25 °C and 23.9 GHz at 80 °C [58].

With the demand for an increased communication rate in short-distance communication systems, the volume of the active area must be reduced to further increase the bandwidth. The width can be reduced using a buried heterogeneous structure or a passive distributed-feedback Bragg reflector (DR), or an active distributed reflector (ADR) can be integrated on one or both sides of the gain region.

In 2018, Huazhong University of Science and Technology made a laser with a 150-μm cavity length using an integrated ADR and conducted a communication-transmission experiment of 28 Gbit/s over 10 km [59]. Optimizing the active area increased the 3-dB bandwidth to approximately 30 dB, but no significant progress

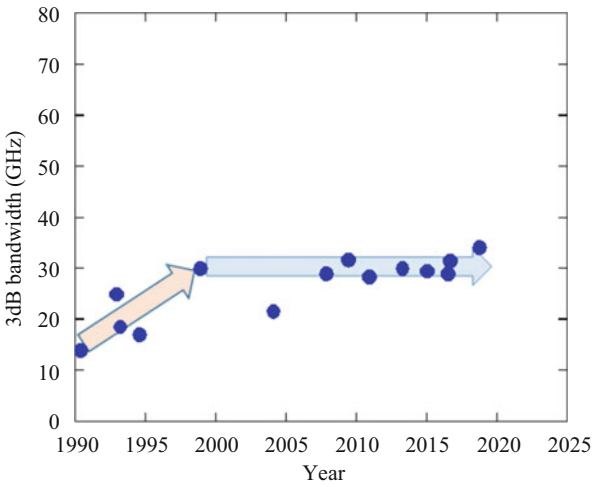


Fig. 4.8 3-dB bandwidth changes of direct-modulation lasers

has been made since then. The change in the 3-dB bandwidth of the direct-modulation laser is shown in Fig. 4.8.

To overcome the bandwidth limitation, researchers have proposed using the photon–photon resonance (PPR) effect produced by optical feedback inside the device. This introduces a response peak at the high frequency of the relaxation oscillation of the laser to broaden its bandwidth [27].

In 2017, Finisar Company achieved a PPR effect through a DR structure and achieved a 3-dB bandwidth of 55 GHz. The 112-Gbit/s PAM4 signal modulation also passed the verification [60]. In 2020, the instrument used by Yanaoka et al. produced a PPR effect near 95 GHz, obtained a 3-dB bandwidth of 108 GHz and a PAM4 signal of 256 Gbit/s, and carried out data transmissions within 2 km [61].

The PPR effect is an important means to significantly increase the bandwidth of the DML; however, it can only can be applied to a DML with external-cavity feedback [62]. Figure 4.9 shows the case of a 3-dB bandwidth in a PPR application. The characteristics of the active region are important factors related to the laser modulation bandwidth, and a large portion of them are related to the relaxation-oscillation frequency. The parasitic parameters of the laser and non-active region factors, such as heat generation, also influence the bandwidth [63].

External Modulation

In external modulation, after the laser is generated, the modulator is placed on the external optical path of the laser, and the physical characteristics of the modulator are changed by modulating the signal. When the laser passes through the modulator, certain parameters of the light wave are modulated and changed [3]. The principle of

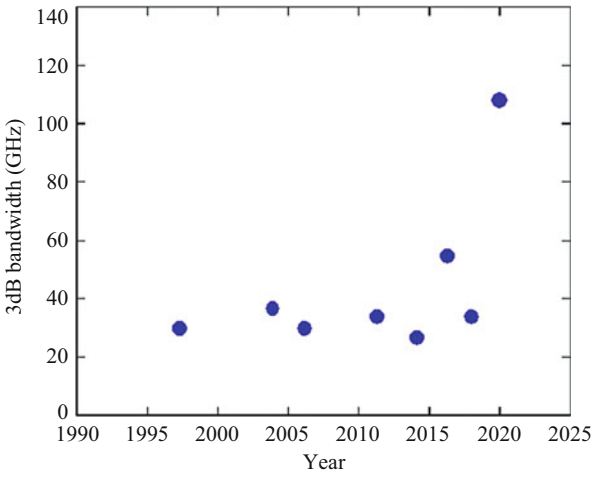


Fig. 4.9 3-dB bandwidth change of a direct-modulation laser when PPR is applied

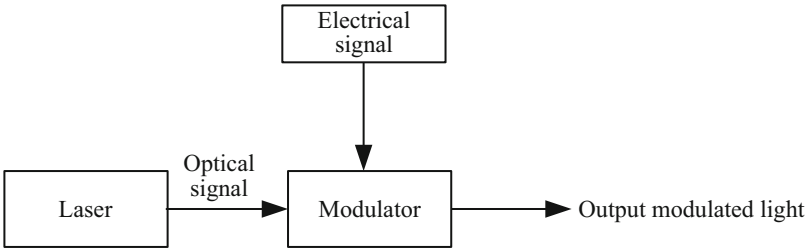


Fig. 4.10 External modulation system [53]

the system is illustrated in Fig. 4.10. External modulation changes the output-laser parameters.

Compared to internal modulation, external modulation applications are more flexible and more easily achieve the required modulation effect. However, because external modulation is based on the existing laser output, it has certain limitations. For example, the laser-peak output cannot be changed by external modulation but can be changed by internal modulation.

Depending on the different working mechanisms, external modulation includes electro-optic, acousto-optic, and magneto-optic modulations. Since 1960, the acousto-optic modulator, based on the acousto-optic effect, Pockels electro-optic modulator, and Faraday rotation magneto-optic modulator have been developed one after another, thus realizing long-distance optical-signal transmissions.

It has been difficult for them to enter the practical stage of laser communications because of the long transit time of acousto-optic modulation and the slow response speed of magneto-optical modulators. Electro-optic modulators meet the requirements of optical information interconnection and transmission and are developing towards high speed, low power consumption, and small volume.

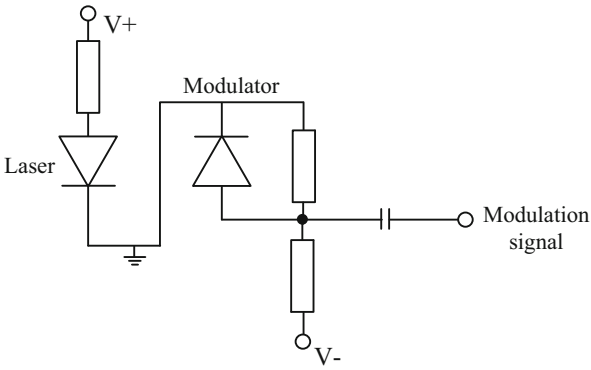


Fig. 4.11 Bias-T structure of an EML

To solve the relaxation-oscillation frequency and frequency-chirp problems when the laser is directly modulated, an electroabsorption modulated laser (EML) has been developed. Generally, a set of biases is applied to the laser, and the bias current is slightly higher than the threshold current. In an EML, the laser is the light-emitting element and the electroabsorption modulator is loaded with information. A bias current is applied to the laser, and the electroabsorption modulator receives a signal-modulation voltage (a bias voltage is also required). The bias-T structure diagram is shown in Fig. 4.11.

The EML, which is monolithically integrated by a distributed-feedback (DFB) laser and an electroabsorption (EA) modulator, is one of the earliest developed and most widely used photonic integrated devices. Its integrated light source has a large modulation bandwidth and low-frequency chirp. Compared with the DML, the EML has advantages in the chirp effect, extinction ratio, eye pattern, jitter, transmission distance, and other performance metrics, and can achieve higher-speed and longer-distance transmissions [64]. The core of the EML is the wavelength matching between the DFB laser and the EA modulator to ensure that the laser output light can pass through the modulator without damage in the zero-modulation bias state.

Research institutes worldwide have conducted studies on high-speed EML-integrated light sources using different integration schemes. For example, the Coldren research group of the University of California, Santa Barbara (UCSB) has developed a high-power integrated light source using selective area etching and quantum-well disorder technology [65]. The Wang Wei Research Group, Institute of Semiconductors, Chinese Academy of Sciences, realized a tunable laser, EA modulator-integrated light source, and array chip using regional selective-epitaxy technology [66]. At present, the modulation bandwidth of single-channel electro-absorption modulation laser technology reaches 59 GHz with a transmission rate of 100 Gbit/s, and it will continue to move towards higher speeds, array types, and lower costs in the future.

In 1970, several countries began to study high-speed $LiNbO_3$ electro-optic modulators. To improve the information-transmission capacity of optical

communications, parameter optimization, thin-film lithium niobate (TFLN), and photon-integration technologies have been proposed [67]. To meet the increasing demand for higher data rates, digital coherent-transmission technology, which is a high-speed coherent signal with a high-order modulation format generated by in-phase quadrature-phase (IQ) modulation, has also received attention.

Parameter Optimization

Traditional $LiNbO_3$ (LN) modulators rely on titanium diffusion or proton-exchange optical waveguides with weak optical limitations; both have large mode sizes and are unsuitable for photonic integrated circuits. Using ridge optical waveguides can effectively reduce the half-wave voltage. The electrode structure changes from an early lumped structure to a traveling-wave electrode structure. The new structure is not affected by the RC time constant of the external circuit, and modulators with bandwidths of hundreds of GHz can be manufactured.

In 1994, Wang et al. made a Ti:$LiNbO_3$ electro-optic modulator with a traveling-wave electrode. This is the earliest reported external modulator with the new electrode structure in China, and its 3-dB modulation bandwidth reaches 8 GHz [68].

In 2001, the Ti:$LiNbO_3$ modulator made by Zhang et al. adopted a T-shaped composite electrode structure, and its bandwidth was increased to 40 GHz. This research achievement was the best in China at that time [69].

In 2005, Gao et al. proposed a combination of a ridge waveguide and T-shaped electrode, which reduced the loss while having a large characteristic impedance. They successfully designed an example of a modulator with a bandwidth of 153 GHz [70].

By 2022, using photolithography, wet etching, high-quality waveguides, and traveling-wave electrodes, Young et al. experimented with a TFLN Mach–Zehnder modulator with a 3-dB bandwidth exceeding 110 GHz and achieved a PAM6 of 250 Gbit/s and PAM4 of 200 Gbit/s [71].

Photonic Integration and Thin-Film Lithium Niobate (TFLN)

To obtain a smaller half-wave voltage and reduce the power consumption of the device, it is hoped that a high refractive-index contrast optical waveguide can be realized by further improving the material and processing technology of single-crystal TFLN. The TFLN material not only retains all the advantages of the traditional LN material but also overcomes all its disadvantages.

In 2018, the TFLN electro-optic modulator based on dry-etching single-crystal integration reported by Loncar's research group in *Nature* had a total length of 1–2 cm, a half-wave voltage as low as 1.4 V, an on-chip optical loss of less than 0.5 dB, and a data-transmission rate as high as 210 Gbit/s [72].

Photonic-integration technology is used to achieve optical communication and interconnection with lower power consumption and higher data transmission. Silicon-based optoelectronic integration has become an important aspect in the development of this technology.

In 2019, Cai et al. completed innovative research on a heterogeneous integrated electro-optic modulating chip of silicon and lithium niobate and retained the

respective advantages of silicon-based and LN materials. The obtained modulation bandwidth exceeded 70 GHz, and its modulation rate (2.2 V·cm) was four times higher than that of a conventional LN modulator. It also showed an ultrahigh data-modulation rate of 112 Gbit/s and a low power consumption of 170 fJ/bit [73].

Based on previous technology, a low-loss TFLN electro-optic modulator with a half-wave voltage of 2.6 V and a modulation bandwidth exceeding 56 GHz was developed. In 2019, an experiment was conducted using this equipment to achieve the highest modulation rate in the world. The experiment verified that, in a single channel, the generated high-speed optical signal reached 120 Gbit/s with NRZ and 220 Gbit/s with PAM4 [73].

In 2021, based on previous research, Professor Loncar's research group created a high-performance modulator with the microwave loss reduced by approximately three times on a TFLN platform of the Shi Ying substrate using a capacitive-loading traveling-wave electrode, which once again broke the upper limit of the bandwidth voltage of TFLN [74].

IQ Modulation

To explore the bandwidth potential of optical-fiber communication and further enhance the transmission capacity of optical communication, around 2008, coherent-optical communication moved from the background to the foreground. In coherent-optical transmission, in addition to amplitude modulation, external modulation methods, such as phase-shift keying (PSK), quadrature phase-shift keying (QPSK), and quadrature amplitude modulation (QAM), can be used to achieve frequency and phase modulation. Owing to the low data-rate requirement, the initial wireless communication adopted analog modulation, such as amplitude modulation (AM), frequency modulation (FM), or phase modulation (PM).

Market demand has not driven progress in large-scale communication technology for a long time. However, with the increasing demand for satellite communication and personal-communication services, traditional analog modulation cannot meet the rate requirement; therefore, it is necessary to seek modulation technology that supports higher data rates. IQ modulation technology is increasingly being used, and almost all wireless communications require this modulation technology.

The main reasons for adopting IQ modulation are as follows:

1. IQ modulation can achieve a higher data rate by increasing the symbol rate or adopting high-order modulation. It is very convenient and flexible, and this is far from the traditional analog modulation.
2. IQ modulation is easier to implement for high-speed communication. IQ modulation can very conveniently map symbols onto vector-coordinate systems, thus completing digital modulation. Similarly, on the receiving side, it can conveniently demodulate the original data bit stream according to the symbol mapping.
3. IQ modulation has a higher spectrum efficiency.

In the last 10 years, in-phase quadrature-phase (IQ) modulators based on LN waveguides with low refractive-index contrast have been the first choice to realize

more modulation formats and have become key devices in digital coherent-optical transmission systems. However, owing to the limitations of the material itself, it is difficult to improve the driving voltage, modulation bandwidth, and integration of LN-IQ modulators, which hinders the upgrading and application of optical-transmission systems.

In 2020, Xu et al. designed high-performance IQ modulator chips for TFLN materials, and their IQ modulator achieved the lowest insertion loss to date. They also generated high-order coherent modulation signals at ultrahigh rates, including 220 Gbit/s per 110 Gbaud and 10 Gbit/s per 80 Gbaud using QPSK.

In 2021, Fujitsu introduced a TFLN modulator and focused on a 130-Gbaud coherent driver module, which achieved a low modulation loss, small package size, and 65-GHz bandwidth. Compared to the InP modulator, the TFLN modulator has a wider bandwidth and lower temperature sensitivity, usually supports broadband operation in the C band, and can eliminate the need for refrigeration, thus reducing the cost and size.

In 2022, Cai et al. first produced a dual-polarization IQ modulator for coherent-optical communication on a TFLN platform, which achieved a CMOS-compatible driving voltage and electro-optical bandwidth of up to 110 GHz. Owing to the excellent bandwidth-voltage performance of the device, a net bit rate of 1.96 Tbit/s per carrier was achieved with an electrical-power consumption of only 1.04 fJ/bit [75]. Coherent-optical communication technology can improve the transmission rate and limit the bandwidth; however, it also results in a high bit-error rate.

Lasers and Modulation Technology

Since 2000, Xi'an University of Technology has conducted extensive basic theoretical research and practical verification, and made progress in optical-communication application fields, such as transmission characteristics in laser atmospheric turbulence, coding and decoding in wireless laser communication systems, optical wireless subcarrier modulation, optical-transceiver system design, capture and alignment-tracking technology, optical orthogonal frequency-division multiplexing modulation, and optical MIMO system research [76].

In 2011, Chen et al. investigated subcarrier-modulation optical wireless communication systems and their key technologies for atmospheric-impact suppression, designed optical wireless communication system hardware for four-way frequency-shift keying (4FSK) and 16PSK modulation modes, and conducted short-distance indoor optical wireless subcarrier-modulation experiments, successfully verifying the feasibility of the system [77].

In 2016, a free-space optical (FSO) communication system terminal with a maximum weight of 8 kg, based on intensity modulation/direct detection (IM/DD), was developed. The terminal was equipped with standard Ethernet data-access capability and safely transmitted duplex data, voice and images.

Laser-Characteristic Optimization

Semiconductor lasers, with high electro-optical conversion efficiency (over 50%), a long service life (over 100,000 h), reliable operation, and easy modulation, emit light waves with a frequency higher than that of microwaves and are widely used. When a signal is loaded onto the LD, nonlinear distortion appears near the threshold and directly affects the system performance [78]. A theoretical derivation and simulation analysis were carried out for its response characteristics, compensation scheme, and proposed power-control system, and a method to improve the nonlinear effect of the laser was obtained.

Response Characteristics

The performance of a communication system depends on its laser-response characteristics. Generally, the response requires a rate equation. The following two formulas briefly explain the relationship between the supply, generation, and disappearance of photons and carriers in a double-heterojunction semiconductor laser (DH-LD) resonator. The rate equation is simplified to [79]:

$$\frac{dN}{dt} = \frac{I}{qV} - A(N - N_{\text{om}})S - \frac{N}{\tau_N} \tag{4.1}$$

$$\frac{dS}{dt} = \Gamma\beta\frac{N}{\tau_N} - \frac{S}{\tau_p} + \Gamma A(N - N_{\text{om}})S, \tag{4.2}$$

where N is the carrier density, S is the photon density, I is the current injected into the gain region, N_{om} is the transparent carrier density, and q is the amount of electron charge. $V = WLD$, where V is the volume of the active region, and W, L, and D are the width, cavity length, and thickness of the gain region, respectively. A is the optical-gain coefficient, τ_p and τ_n are the photon lifetime and carrier lifetime in the resonant cavity, respectively, Γ is the optical-confinement factor, and β is the spontaneous emission.

Modulation-Response Characteristics

The modulation-response characteristics directly reflect the high-speed characteristics of the laser in the frequency domain, and the nonlinearity of the laser reflects the nonlinearity of the modulation response. As shown in Fig. 4.12, the high-frequency modulation characteristics of the laser are closely related to the relaxation-oscillation frequency f_0. Increasing f_0 can increase the modulation frequency ω of the laser. When the resonance peak shifts to the right, the low-frequency band that can be modulated is expanded.

Relaxation oscillation means that when current is injected into the laser, the number of particle inversions oscillates near the threshold; that is, the number of particle inversions is lower than the threshold when stimulated radiation occurs but higher than the threshold when pumped. This phenomenon results in a series of short

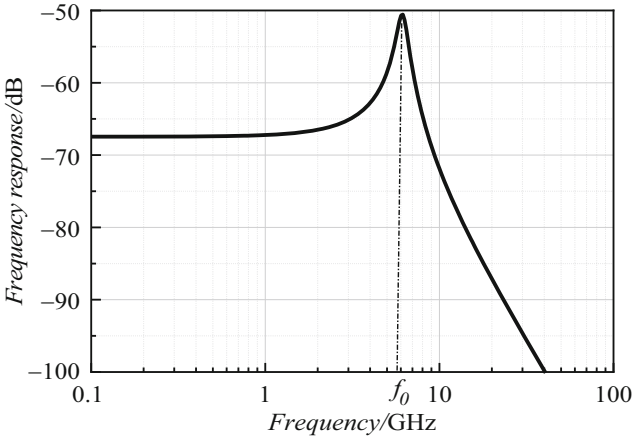


Fig. 4.12 DH-LD frequency-response curve [80]

pulse sequences instead of a smooth pulse. Therefore, the output power and carrier concentration will fluctuate between a steady state and an unsteady state, which is called the relaxation-oscillation phenomenon [81].

The dynamic distortion related to the frequency in nonlinear distortion is caused by the relaxation oscillation. Under small-signal modulation, the relaxation-oscillation frequency can be obtained as:

$$f_0 = \omega = \left(\frac{a_g S_0}{\tau_p (1 + \varepsilon S_0)} \right)^{1/2}. \tag{4.3}$$

It can be seen from the above formula that the relaxation-oscillation frequency f_0 can be increased by increasing the differential-gain coefficient a_g and steady-state photon density S_0, and decreasing the photon lifetime τ_p and gain-saturation factor ε. Subsequently, we simulated the model to verify this correlation.

Figure 4.13 shows the influence of changing the related relaxation-oscillation parameters on the modulation-response characteristics. The simulation results were consistent with the theoretical conclusions.

The modulation-response characteristics are not only related to the internal relaxation-oscillation parameters of the laser but also to the external parasitic parameters. As shown in Fig. 4.14, to achieve a high frequency response, the values of the parasitic parameters can be reasonably regulated by changing the laser packaging type, and the series resistance and parallel capacitance can be reduced.

Impulse-Response Characteristics

The impulse response refers to the response characteristics in the time domain (time amplitude) obtained by a system under test when it receives a pulse-excitation signal

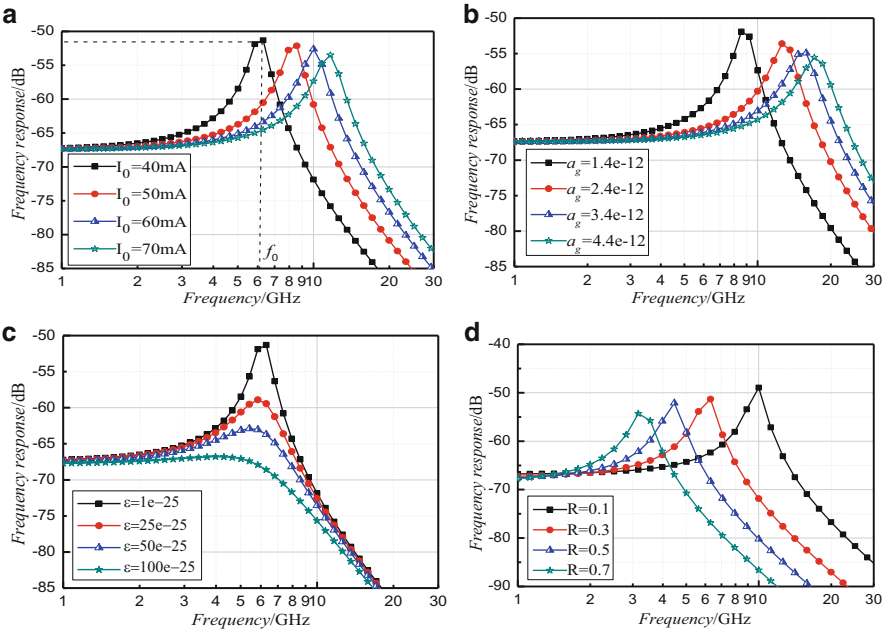


Fig. 4.13 Frequency-response curves under the influence of different relaxation-oscillation parameters [2]. (**a**) Different bias currents I_0. (**b**) Different differential-gain coefficients a_g. (**c**) Different gain-saturation factors ε. (**d**) Different reflectivities R

input. Figure 4.15 shows that relaxation oscillation also occurs before reaching a steady state. We used the transient analysis from the PSpice simulation environment to verify the simulation.

As shown in Fig. 4.16, the trends in the time and frequency domains are the same. When I_0 and a_g increase, ε and R decrease. This increases the relaxation-oscillation frequency f_0, shortens the oscillation process, and improves the waveform distortion (Figs. 4.17 and 4.18).

Compensation Scheme

Studying the nonlinear characteristics of a laser and its nonlinear-distortion compensation methods can effectively improve the nonlinearity produced by the laser. Commonly used compensation methods include optical (electrical) negative feedback, positive feedback, predistortion, etc.

In different turbulence-intensity channels, optical wireless multichannel subcarrier binary phase-shift keying (BPSK) modulation was adopted, and the relationship between the modulation index, nonlinear coefficient, and carrier-intermodulation noise ratio (CINR) of a wireless laser communication system was investigated. Furthermore, the selection of the modulation index and nonlinear coefficient to reduce the adverse interference of the nonlinear intermodulation distortion of the laser on the CINR and bit-error rate (BER) was analyzed.

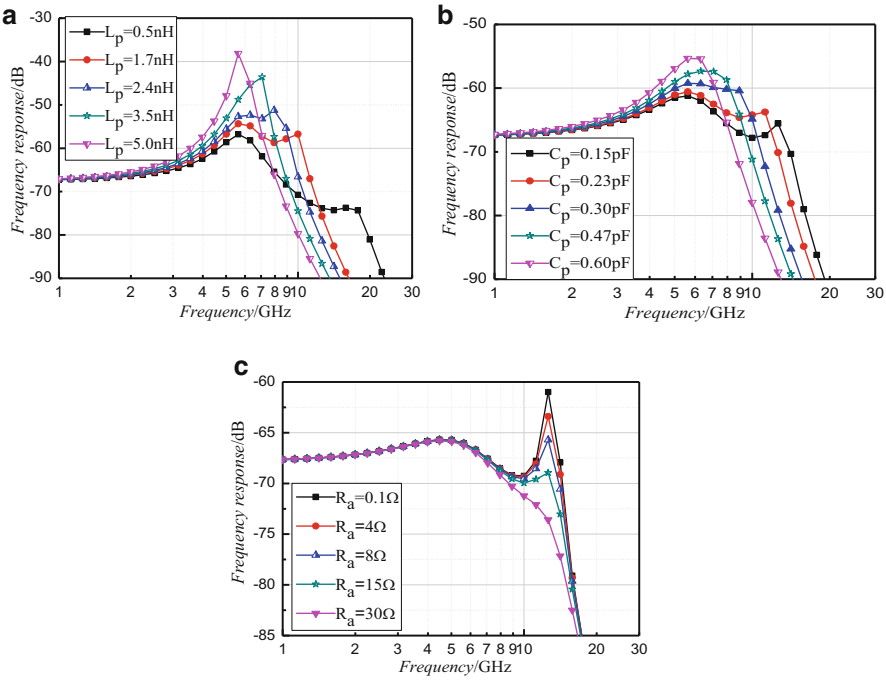


Fig. 4.14 DH-LD frequency-response curves affected by different parasitic parameters [2]. (**a**) Parasitic inductance L_p. (**b**) Parasitic capacitance C_p. (**c**) Parasitic resistance R_a

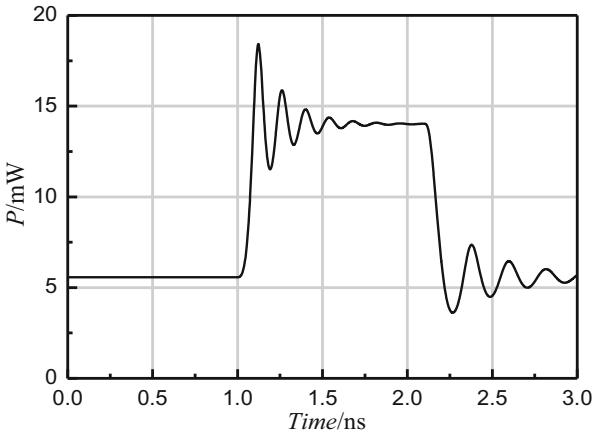


Fig. 4.15 DH-LD impulse-response curve

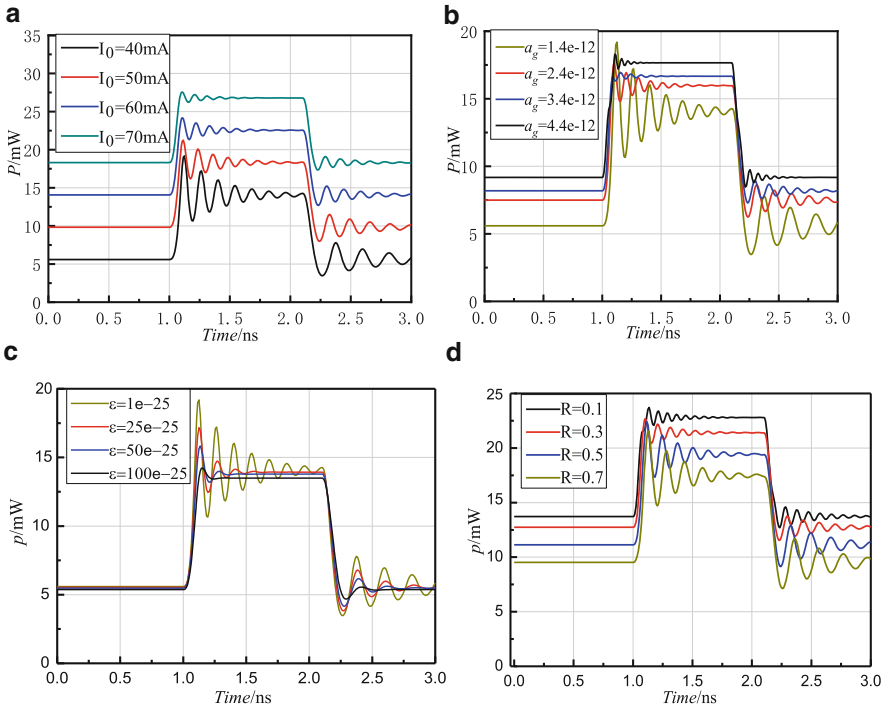


Fig. 4.16 Impulse-response curves under the influence of different relaxation-oscillation parameters [2]. (**a**) Different bias currents I_0. (**b**) Different differential-gain coefficients a_g. (**c**) Different gain-saturation factors ε. (**d**) Different reflectivities R

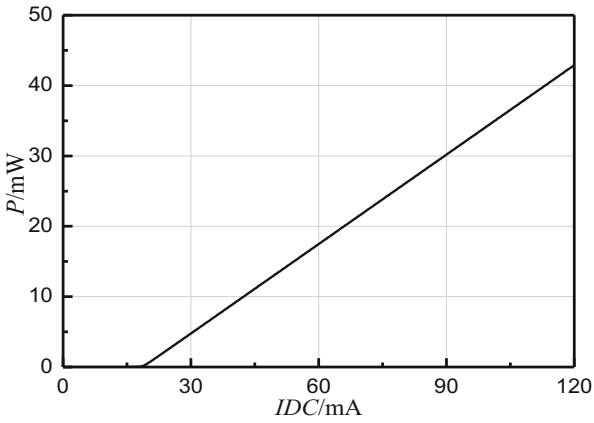


Fig. 4.17 P-I characteristic curves of a DH-LD [2]

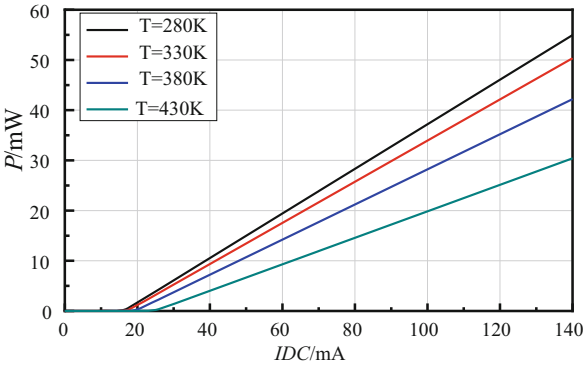


Fig. 4.18 P-I characteristic curves of different temperatures [2]

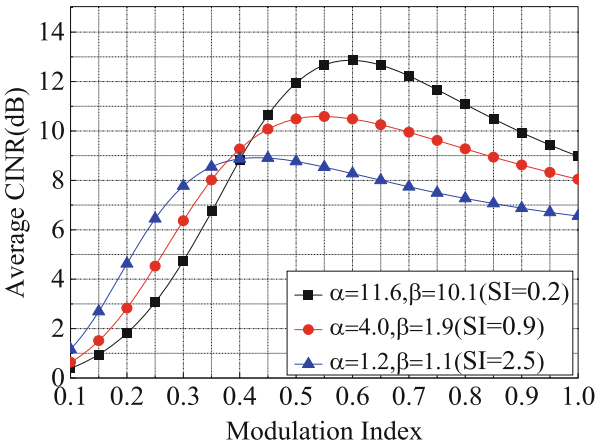


Fig. 4.19 CINR study under a turbulent flow [82]

As shown in Fig. 4.19, under the same turbulence conditions, the closer the modulation index is to the optimal value, the greater the CINR value. When the scintillation index (SI) increases, the optimal modulation index should decrease, and the maximum CINR can be obtained.

Figures 4.20 and 4.21 show that the CINR decreases when the carrier number N increases. When the turbulence intensity is the same, the larger the nonlinear coefficient, the smaller the CINR.

From Fig. 4.22, we can see that when the modulation index is larger than the optimal value, the BER increases with an increase in SI. If the modulation index further increases, the BER becomes saturated and stable under the same modulation index.

A Volterra-series analysis was used to describe the influence of the static nonlinearity of the laser on the non-limiting QPSK-like sinusoidal modulation. The influence of the LD nonlinearity on the signal was simulated by fitting the laser

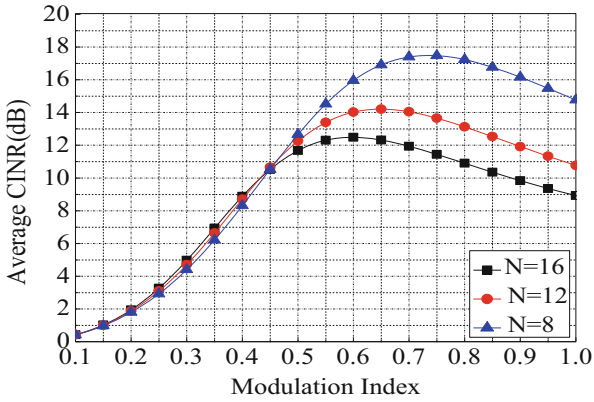


Fig. 4.20 CINR under various carrier numbers [82]

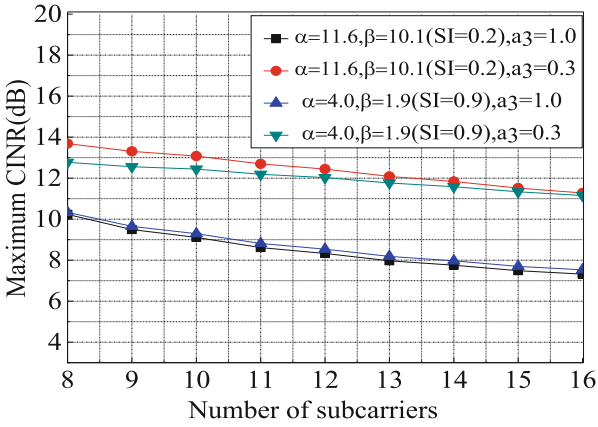


Fig. 4.21 Maximum CINR under different nonlinear coefficients [82]

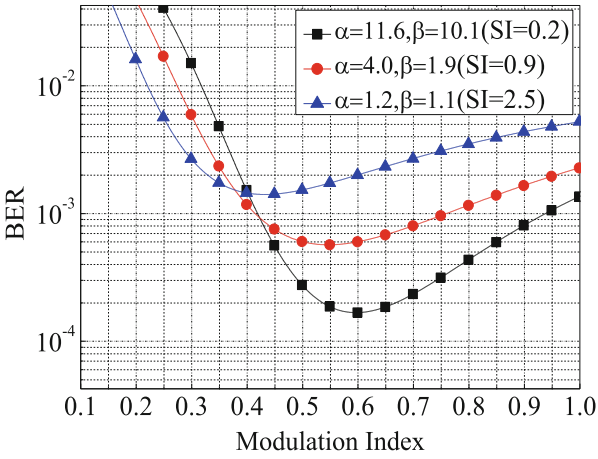


Fig. 4.22 BER under atmospheric turbulence [82]

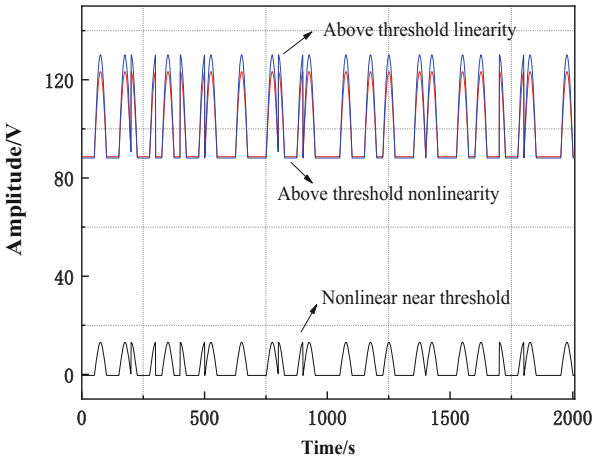


Fig. 4.23 Time-domain waveform directly after LD and predistortion [82]

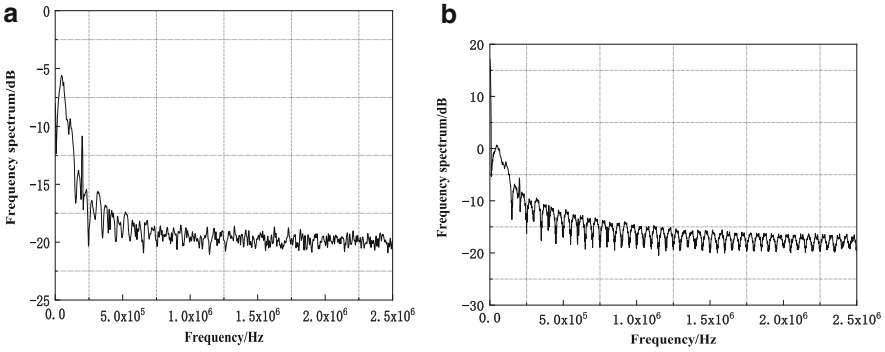


Fig. 4.24 Spectrum before and after predistortion compensation [83]. (**a**) Before predistortion. (**b**) After predistortion

characteristic curve, and the best predistortion method was selected for nonlinear compensation. Through simulation, the time-domain waveform, frequency spectrum, power spectrum, and constellation of the non-limiting QPSK sinusoidal-modulation signal before and after predistortion compensation were obtained, as shown in Figs. 4.23, 4.24, 4.25, and 4.26.

Power Control System

In optical wireless communication systems, it is necessary for an optical transmitter to provide a stable and reliable light source to ensure the communication quality. The ability of the receiver to receive a high-quality signal is closely related to the power stability of the transmitter. Therefore, Xi'an University of Technology adopted the principle of photoelectric feedback to realize a power control system for a semiconductor laser to ensure the power stability of the laser light source.

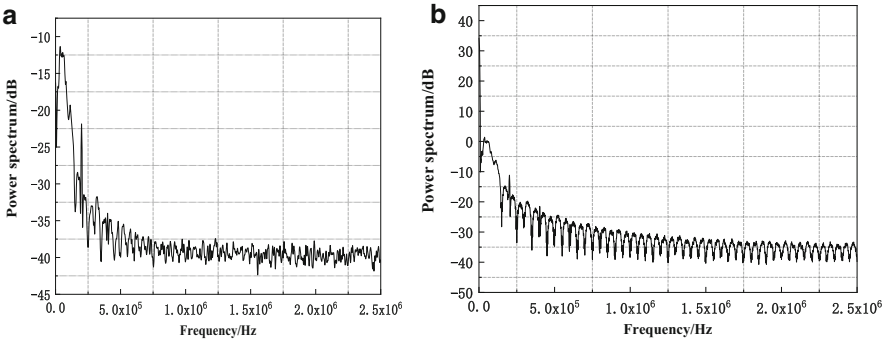


Fig. 4.25 Power spectrum before and after predistortion compensation [83]. (**a**) Before predistortion. (**b**) After predistortion

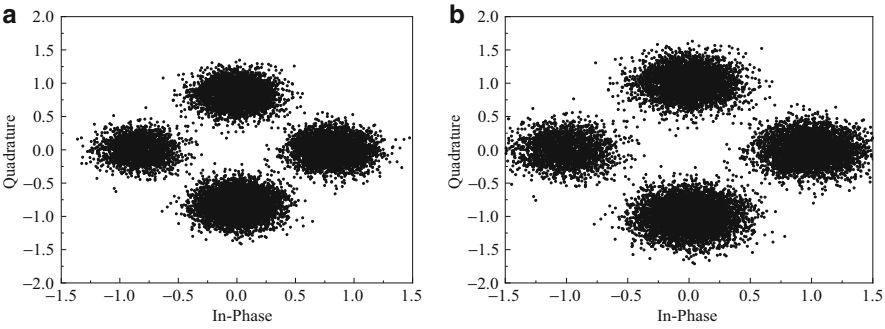


Fig. 4.26 Constellation before and after predistortion compensation [83]. (**a**) Before predistortion. (**b**) After predistortion

The system consists of an automatic power control (APC) and automatic temperature control (ATC). The APC includes a constant current source module and a photoelectric feedback module, whereas the ATC includes a temperature-detection module and refrigeration drive module. Figure 4.27 shows the overall block diagram of the system.

The APC circuit was maintained in an open-loop state, and the stability of the semiconductor laser driven by a constant current source was verified. As shown in Fig. 4.28, the driving performance of the constant current source was relatively stable in the experiment, and the stability was better than 0.014%.

The APC circuit was tested and verified at room temperature; the results are shown in Fig. 4.29. When the circuit is an open loop, the output optical power decreases with the extension of the working time, and the fluctuation range of the output optical power is larger, with a power stability of 4.1%. Under a closed-circuit condition, the output optical power tends to be stable as a whole, and there is no

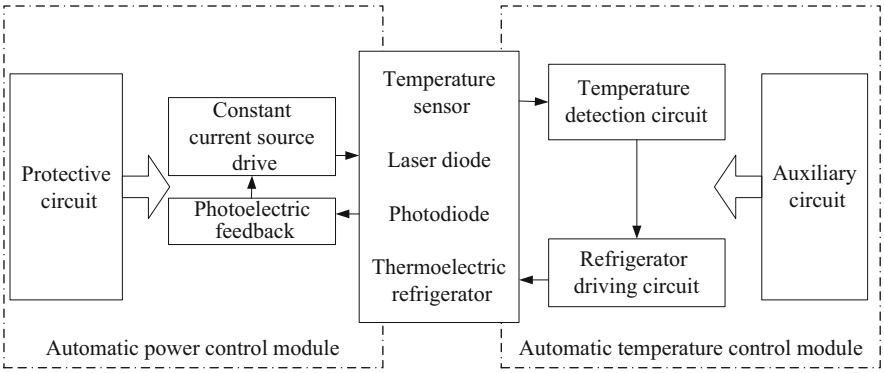


Fig. 4.27 Overall block diagram of the system [84]

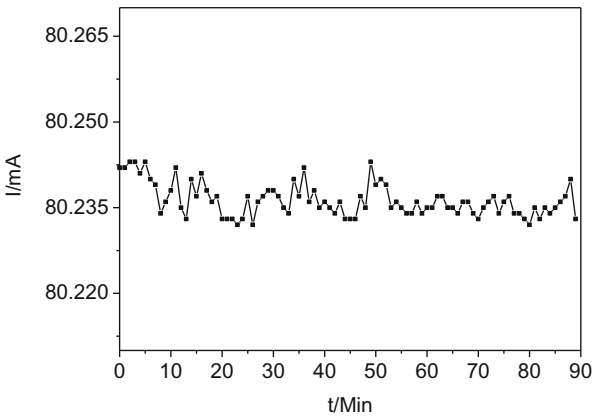


Fig. 4.28 Changes in the driving current under a constant current source [84]

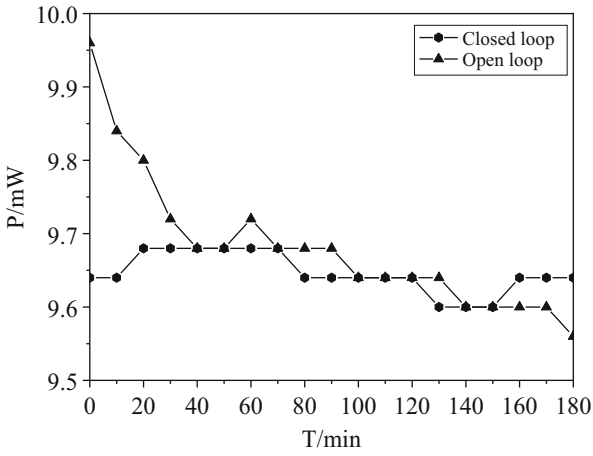


Fig. 4.29 APC output optical power in open and closed loops [84]

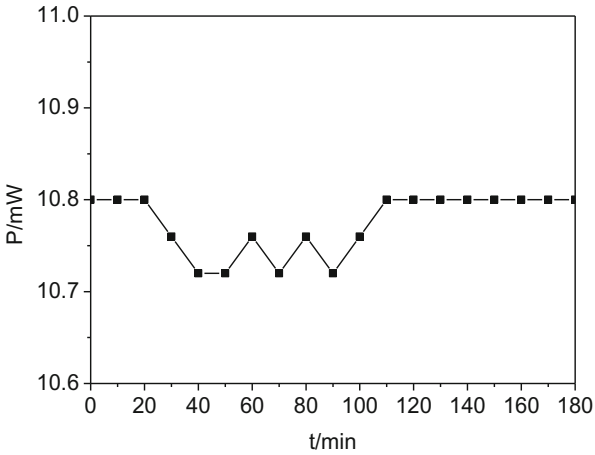


Fig. 4.30 Output power with APC and ATC working simultaneously [84]

obvious fluctuation with the extension of the working time. The stability is calculated to be 0.83%.

Therefore, when the circuit loop is closed, that is, when the reverse input of the proportional amplifier is connected to the photoelectric feedback module, it can stabilize the output optical power, thus verifying the feasibility and effectiveness of the automatic power-control circuit.

Figure 4.30 shows the fluctuation chart of the output power with working time when the APC and ATC are working simultaneously. During the entire operation process, the output power fluctuates between 11.7 MW and 11.8 MW, and the entire process is relatively flat, which proves that the ATC circuit can ensure that the output power is more stable; the calculated stability is better than 0.74%.

The ATC circuit can provide a stable ambient temperature during the semiconductor-laser operation. As shown in Fig. 4.31, when the ATC circuit works, the ambient temperature of the semiconductor laser is kept around 27 °C, and the fluctuation is less than 0.1 °C; thus, the stability is 0.74%.

Modulation Experiment

Xi'an University of Technology systematically conducted experiments on laser modulation, pulse-position modulation, and multi-carrier modulation, and improved the system structure and inevitable nonlinear effects.

Internal Modulation

To compare the first three generations of atmospheric laser communication machines, Xi'an University of Technology investigated a high-speed internal-

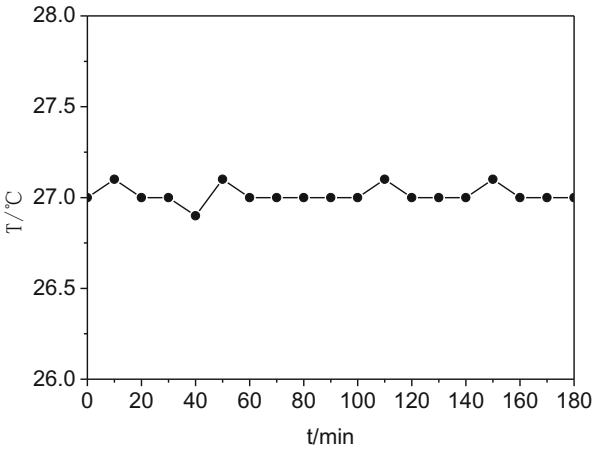


Fig. 4.31 Temperature-change curve with ATC working [84]

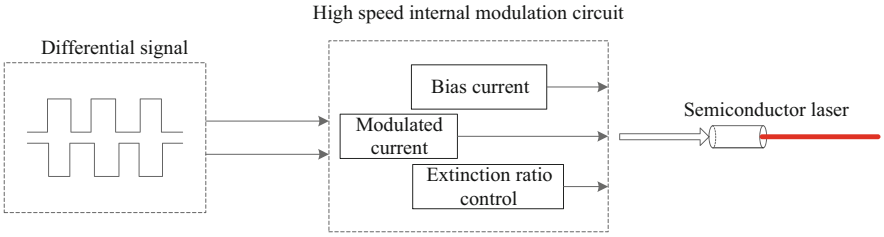


Fig. 4.32 Block diagram of a high-speed internal-modulation circuit [53]

modulation system. Figure 4.32 shows a block diagram of the internal-modulation circuit structure. After the circuit performance is verified, the entire communication system will be debugged. Currently, the communication requirements for a 100-MHz debugging rate can be realized.

Progress on high-speed direct-modulation lasers in China is slow. At present, companies such as Shaanxi Yuanjie, Wuhan Guangxun, and Ningbo Yuanxin have developed direct-modulation lasers with bandwidths above 20 GHz; therefore, great development prospects still exist.

Electro-optic Modulation

Based on internal modulation, an electro-optic modulation experiment was conducted, as shown in Fig. 4.33. The light emitted by the laser passes through the polarizer and becomes linearly polarized. The light is vertically incident on the LN crystal to which the electric field is applied, and the refractive index and polarization state of the light wave are changed. It is output as an elliptically polarized light.

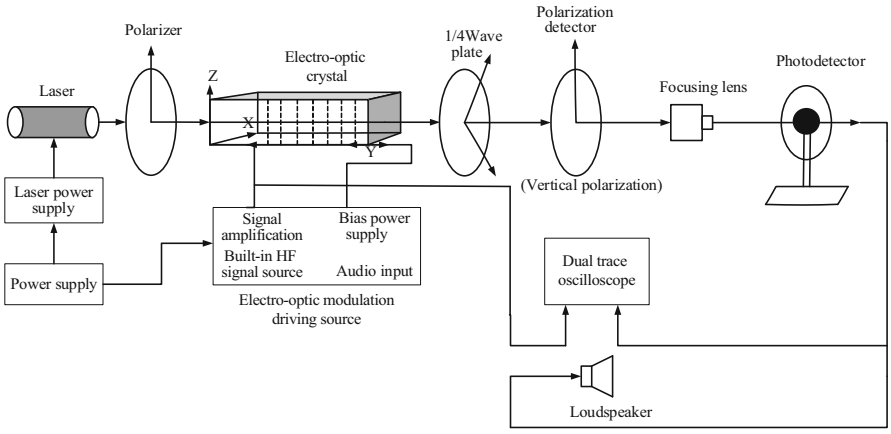


Fig. 4.33 Structural diagram of the electro-optic modulation experimental system [87]

The light can be biased by the DC generated by the modulator driving source or the outgoing light passing through the quarter-wave plate so that the electro-optic modulator can work in the linear region [85]. Elliptically polarized light passes through the analyzer and is output as linearly polarized light, and the light intensity is modulated. Finally, a photoelectric detector receives the modulated light signal, which is observed using an oscilloscope after the photoelectric conversion.

According to the theory of refractive-index ellipsoids, there is a relationship between the refractive index of the electro-optic crystal, output light intensity, transmittance of the light modulator, and temperature [86].

Influence of the Temperature on the Refractive Index of an Electro-optic Crystal

The LN crystal is sensitive to changes in external temperature, and the refractive index of the electro-optic crystal has the greatest influence. Figure 4.34 shows that the refractive index changes linearly with the temperature, and changes in ambient temperature have a greater influence on the refractive index of light.

Influence of the Temperature Electro-optic Modulator on the Output Light Intensity

Depending on whether the direction of the applied electric field is parallel or perpendicular to the light-wave propagation direction, the electro-optic modulation can be divided into longitudinal and transverse. If the intensity of the incident light is I_0, the intensity I_1 of the outgoing light is $I_1 = I_0 \sin^2\left(\frac{\delta}{2}\right)$, where δ is the phase delay.

As shown in Fig. 4.35, when the input light intensity, applied voltage, and light wavelength are fixed, the output light intensity increases linearly with the increasing temperature during longitudinal modulation, which is related to the change in the crystal refractive index.

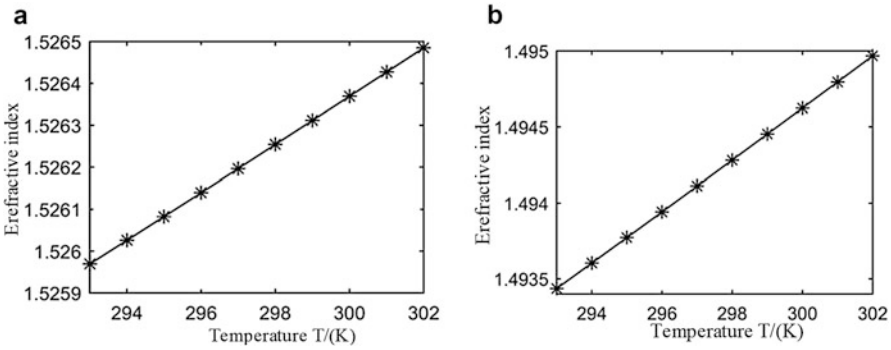


Fig. 4.34 Trends of the refractive index of O and E light with temperature [87]. (**a**) *O* light. (**b**) *E light*

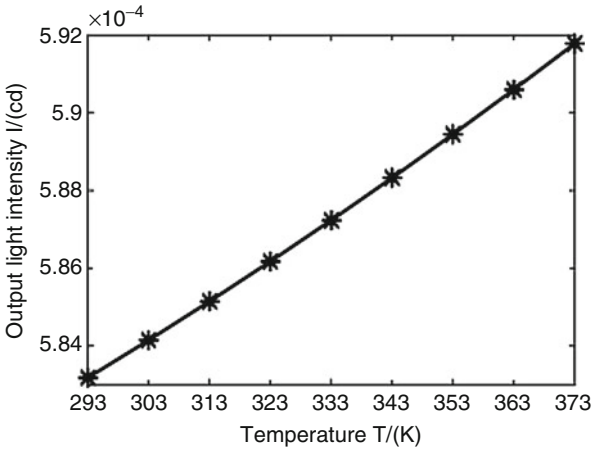


Fig. 4.35 Variation trend of longitudinal-modulated output light intensity with temperature [87]

Under transverse electro-optic modulation, the change in the output light intensity with the temperature is related to many factors, such as the crystal refractive index, linear-expansion coefficient, initial crystal length L_0, and thickness d_0. Figure 4.36 shows the influence of the temperature change on the output light intensity when the crystal size and linear expansion coefficient are constant. In practice, transverse electro-optic modulation is more flexible, and the half-wave voltage can be reduced by reducing the thickness of the modulation device or increasing its length; however, the processing requirements are relatively high.

Influence of the Temperature on the Light-Modulator Transmittance

The light transmittance can be expressed as $\frac{I_1}{I_0} = \sin^2\left(\frac{\delta}{2}\right)$; therefore, the influence of the temperature on the transmittance of the light modulator is its influence on the

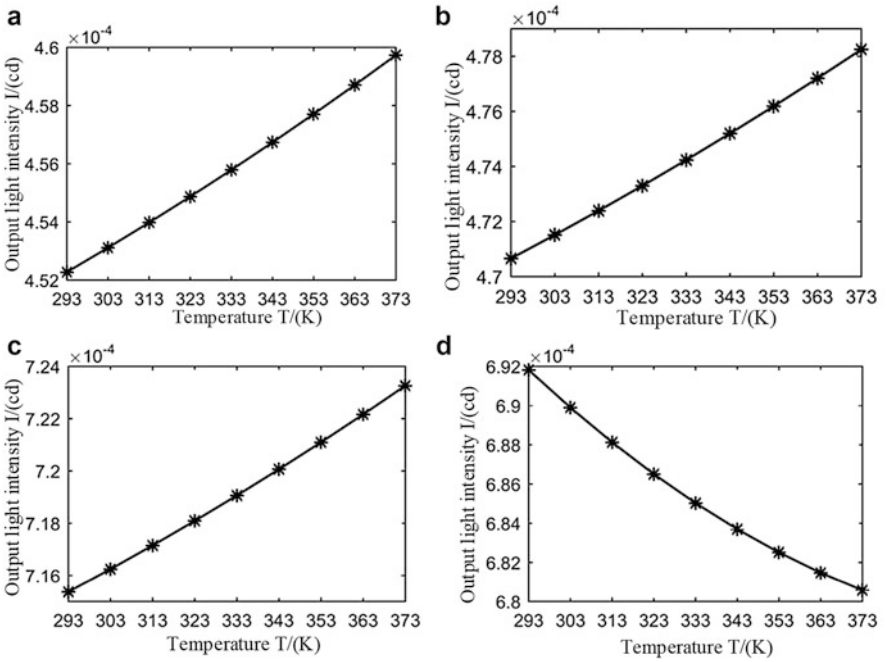


Fig. 4.36 Curves of the relationship between the transverse-modulation output light intensity and temperature change [87]. (**a**) $L_{01} = 0.044$ m, $d_{01} = 0.041$ m. (**b**) $L_{01} = 0.044$ m, $d_{01} = 0.04$ m. (**c**) $L_{01} = 0.044$ m, $d_{01} = 0.03$ m. (**d**) $L_{01} = 0.06$ m, $d_{01} = 0.01$ m

light retardation. The temperature influence on the retardation of horizontal and vertical modulation is different.

In transverse electro-optic modulation, the crystal retardation is affected by multiple factors. The following can be determined from Eq. (4.4): With a given external voltage and 1000-V DC voltage, when the temperature changes by one degree, the delay will correspondingly change by $1.325\Pi \times 10^{-5}$ radians, which is about 0.002°. At a given temperature, when the voltage changes by one volt, the delay changes by $1.3187\Pi \times 10^{-3}$ radians, which is about 0.237°.

$$\delta_T = \left[1.3187\pi \times 10^{-3} + 1.325 \times 10^{-8} \cdot (T - T_0)\right] \cdot U, \tag{4.4}$$

where T_0 is the temperature of the initial state and U is the applied voltage.

As shown in Figs. 4.34, 4.35, and 4.36, the external temperature significantly influences the electro-optic modulator, which causes the linear operating point of the modulator to shift. The designed electro-optic-modulation driving source was used in the electro-optic communication experiment, and the effectiveness of the driving source was verified. The effectiveness was not only reflected in ensuring the normal operation of the electro-optic modulator and stable linear modulation of the laser, to

a certain extent, but also in the improvement of the modulation rate, which increased to 3 Mbit/s after testing.

Summary

Signal generation cannot be separated from suitable light sources. Not all light sources can be used in optical communication. The light sources used in optical communication must satisfy a series of requirements, such as adequate wavelength, power, spectral linewidth, modulation bandwidth, integration, and cost. Currently, almost all light sources used in optical communications are fabricated using semiconductor technology.

The integration of new physics, concepts, and technologies with semiconductor lasers has introduced fresh blood into their development. Many new lasers have been developed through cross-penetration using optics, electromagnetism, microelectronics, topology, and quantum mechanics. They either have large-scale integrated application prospects, excellent beam and spectrum quality, or a higher and more stable output power. The development of these new laser systems represents an advanced level of semiconductor laser technology and also reflects the development status of physical theory, engineering technology, and preparation technology, which are worthy of further study.

1. Optically pumped vertical-external-cavity surface-emitting laser

Optically pumped vertical-external-cavity surface-emitting lasers (OP-VECSELs), also known as optically pumped semiconductor lasers (OPSLs) or semiconductor disk lasers (SDLs), are a combination of semiconductor and solid-state lasers. The gain chip is made of a semiconductor material, which is very similar to a vertical-cavity surface-emitting laser (VCSEL). The resonant cavity adopts the configuration of a solid-state laser, which is typically composed of a distributed Bragg reflector (DBR) on a semiconductor chip and an external cavity mirror. The pumping mode typically involves optical pumping, which provides a more flexible working mode and better device performance.

A VECSEL uses a semiconductor chip as a gain substance, which provides multiple wavelength selections and wide-spectrum tuning ranges. The optical cavity based on a solid-state laser makes it easy to insert optical elements into the cavity, such as pulse compression, sum frequency, difference frequency, and beam shaping. These can produce ultrashort pulse lasers, special wavelength lasers, terahertz lasers, and multicolor lasers to meet various special application requirements.

2. Micro-nano laser

Micro-nano lasers usually refer to lasers whose size or mode size is close to or smaller than the wavelength of the emitted light. Their compact structure, low threshold, and low power consumption provide broad application prospects for high-speed modulation. The micro-nano laser is an important part of the future of

integrated optical circuits, optical memory chips, and photonic computers. It is also widely used in the fields of biochips and laser medicine, and has potential application value in wearable devices and other fields.

3. Topological insulator laser

Topological-insulator lasers combine semiconductor laser technology with the concept of a topological insulator in condensed-matter physics. Using the topological concept of an edge state, this laser is insensitive to disturbances and defects in the internal structure of the device, and lasers with high output power, high robustness, and a stable mode can be easily realized. Especially used in high-power lasers and emerging nano-photon lasers, this type of laser is insensitive to scattering loss and random manufacturing defects, making it suitable for high-power mode-locked laser arrays and quantum-information generation and transmission.

References

1. Xizheng Ke, Ke Dong, Optical Wireless Communication [M], Springer Press, 2022,:10–14.
2. Liu Zhaohui. Study on response characteristics of semiconductor laser and photodetector [D]. Xi'an: Xi'an University of Technology, 20110.
3. Xia Min. Principle and Technology of Laser [M]. Beijing: Science Press, 2018.
4. Maiman T H. Stimulated Optical Radiation in Ruby [J]. Essentials of Lasers, 1969, 187(4736): 493–494.
5. Wu Chengjiu. Progress and potential of laser crystals and solid-state lasers [J]. Optoelectronic Technology and Information, 1999, 12(06): 11–16.
6. Bo Y, Geng A, Bi Y, et al. High-Power and High-Quality, Green-Beam Generation by Employing a Thermally Near-Unstable Resonator Design [J]. Applied Optics, 2006, 45(11): 2499–2503.
7. Cheng X, Cui Q, Zhou Y, et al. High Power and High Beam Quality CW Green Beam Generated by Diode-side-pumped Intracavity Frequency Doubled Nd:YAG Laser [J]. Optics Communications, 2009, 282(21): 4288–4291.
8. Singh A, Mukhopadhyay P K, Sharma S K, et al. 82 W Continuous-Wave Green Beam Generation by Intracavity Frequency Doubling of Diode-Side-Pumped Nd:YAG Laser [J]. IEEE Journal of Quantum Electronics, 2011, 47(03): 398–405.
9. Ji B, Zheng X, Cai Z, et al. Compact High Conversion Efficiency Nd:YAG/LBO Green Laser Using Un-Stable V Cavity [J]. Laser Physics, 2012, 22(02): 406–411.
10. Zhao Zhibin, Cheng Cheng, Li Quan, et al. Study on all-solid-state CW 228nm deep ultraviolet laser [J]. Acta Photonica Sinica, 2022, 51(9): 0914003.1-0914003.8.
11. Lin L, Ponnappan R. Heat Transfer Characteristics of Spray Cooling in a Closed Loop [J]. International Journal of Heat & Mass Transfer, 2003, 46(20): 3737–3746.
12. Tzuk Y, Tal A, Goldring S, et al. Diamond Cooling of High Power Diode Pumped Solid State Lasers [J]. IEEE Journal of Quantum Electronics, 2003, 40(03): 262–2610.
13. Li Bingbin, Guo Zhen, Song Xiaolu, et al. Direct heat dissipation from the end-pumped solid-state laser [J]. China Laser, 2009, 36(01): 59–65.
14. Zhang Kuo, Yu Deyang, He Yang, et al. Analysis of heat dissipation performance of compact mid-infrared solid-state laser [J]. China Laser 44(03): 26–31.
15. Gan Qijun, Jiang Benxue, Zhang Pande, et al. Research progress of high average power solid-state lasers [J]. Advances in Laser and Optoelectronics, 2017, 54(01): 29–310.
16. Ke Xizheng,. Introduction to Wireless Laser Communication [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2004.

17. Patel C, Tien P, Mcfee J. CW HIGH-POWER CO_2-N_2-He LASER [J]. Applied Physics Letters, 1965, 7(11): 290–292.
18. Yang Dalin, He Yaokun, Shi Shunsen. Acousto-optic Q-switched waveguide CO2 laser [J]. Laser Technology, 1991, 15(01): 9–14.
19. Xie Jijiang, Pan Qikun, Li Dianjun, et al. Theoretical calculation and experimental study of acousto-optic Q-switched CO2 laser [J]. Chinese lasers ,2011, 38(02): 36–42.
20. Xie Jijiang, Li Dianjun, Zhang Chuansheng, et al. Acousto-optic Q-switched CO2 laser [J]. Optical Precision Engineering, 2009,17 (05): 1008–1014.
21. Pan Qikun, Xie Jijiang, Ruan Peng, et al. Output characteristics of acousto-optic Q-switched CO2 laser [J]. China Optics, 2012, 5(03): 283–289.
22. Pan Qikun, Yu Hanghang, Chen Fei, et al. Experimental study of acousto-optic deflection fast-tuning pulsed CO2 laser [J] .. 12(02): 355–361.
23. Wang Lijun, Ning Yongqiang, Qin Li, et al. Research progress of high power semiconductor lasers [J]. Journal of Luminescence, 2015, 36(01): 1–110.
24. Kogelnik H, Shank C. Stimulated Emission in a Periodic Structure [J]. Applied Physics Letters, 1971, 18(4): 152–154.
25. Kogelnik H, Shank C. Coupled-Wave Theory of Distributed Feedback Lasers [J]. Journal of Applied Physics, 1972, 43(05): 2327–2335.
26. Nakamura M, Yariv A, Yen H. Optically Pumped GaAs Surface Laser with Corrugation Feedback [J]. Applied Physics Letters, 1973, 22(10): 515–516.
27. Lu Dan, Yang Qiulu, Wang Hao, et al. Semiconductor distributed feedback laser in communication band [J]. China Laser, 2020, 47(07): 11–210.
28. Sato K, Yasue S, Yamada K, et al. Room-temperature operation of AlGaN ultraviolet-B laser diode at 298 nm on lattice-relaxed Al0.6Ga0.4N/AlN/sapphire [J]. 2020, 13(3): 031004.1–031004.5.
29. Wang Qiming. Breakthroughs and development of semiconductor lasers in China (invited papers) [J]. China Laser, 2010, 37(09): 2190–2198.
30. Soda H, Iga K, Kitahara C, et al. GaInAsP/InP surface Emitting Injection Lasers [J]. Japanese Journal of Applied Physics, 1979, 18(12): 2329–2330
31. Koyama F, Kinoshita S, Iga K. Room-Temperature Continuous Wave Lasing Characteristics of a GaAs Vertical Cavity Surface-Emitting Laser [J]. Applied Physics Letters, 1989, 55(03): 221–222.
32. Johnson R H, Serkland D K. 17G Directly Modulated Datacom VCSELs: 2008 Conference on Lasers and Electro-Optics and 2008 Conference on Quantum Electronics and Laser Science, San Jose, CA, March 4–9, 2008[C]. Piscataway, NJ: IEEE, 2009.
33. Kuchta D M, Rylyakov A V, Doany F E, et al. A 71-Gb/s NRZ Modulated 850-nm VCSEL-Based Optical Link [J]. IEEE Photonics Technology Letters, 2015, 27(06): 577–580.
34. Kuchta D M, Pepeljugoski P, Kwark Y. VCSEL Modulation at 20 Gb/s over 200 m of Multimode Fiber Using a 3.3 V SiGe Laser Driver IC: 2001 Digest of LEOS Summer Topical Meetings: Advanced Semiconductor Lasers and Applications/Ultraviolet and Blue Lasers and Their Applications/Ultralong Haul DWDM Transmission and Networking/WDM Compo, Copper Mountain, CO, July 30-August 1, 2001[C]. Piscataway, NJ: IEEE, 2001.
35. Westbergh P, Gustavsson J S, Koegel B, et al. 40 Gbit/s Error-Free Operation of Oxide-Confined 850 nm VCSEL [J]. Electronics Letters, 2010, 46(14): 1014–1016.
36. Westbergh P, Haglund E P, Haglund E, et al. High-Speed 850 nm VCSELs Operating Error Free Up to 57 Gbit/s [J]. Electronics Letters, 2013, 49(16): 1021–1022.
37. Kuchta D, Rylyakov A V, Schow C L, et al. 64Gb/s Transmission over 57m MMF Using an NRZ Modulated 850nm VCSEL: Optical Fiber Communication (OFC) Conference 2014, San Francisco, CA, March 9–13, 2014[C]. Piscataway, NJ: IEEE, 2016.
38. Lavrencik J, Varughese S, Ledentsov N, et al. 168Gbps PAM-4 Multimode Fiber Transmission through 50m using 28GHz 850nm Multimode VCSELs: 2020 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, March 8–12, 2020[C]. Piscataway, NJ: IEEE, 2020.

39. Zuo T, Zhang T, Zhang S, et al. 850-nm VCSEL-Based Single-Lane 200-Gbps PAM-4 Transmission for Datacenter Intra-Connections [J]. IEEE Photonics Technology Letters, 2021, 33(18): 1042–1045.
40. Chen Lianghui, Yang Guowen, Liu Yuxian. Research progress of semiconductor lasers [J]. China Laser, 2020, 47(05): 13–31.
41. Liu Lijie, Wu Yuanda, Blair, et al. Research on 1550 nm vertical cavity surface emitting laser for long-distance communication [J]. Journal of Infrared and Millimeter Wave, 2020, 39(04): 397–400.
42. Caliman A, Mereuta A, Suruceanu G, et al. 8 mW Fundamental Mode Output of Wafer-Fused VCSELs Emitting in the 1550-nm Band [J]. Optics Express, 2011, 19(18): 16996–17001.
43. Boehm G, Ortsiefer M, Shau R, et al. Inp-Based VCSEL Technology Covering the Wavelength Range from 1.3 to 2.0μm [J]. Journal of Crystal Growth, 2003, 251(1/4): 748–753.
44. Hofmann W, M Müller, Nadtochiy A, et al. 22-Gb/s Long Wavelength VCSELs [J]. Optics Express, 2009, 17(20): 17547–17554.
45. Han Saiyi, Tian Sicong, Xu Hanyang, et al. Research progress of high-speed 1550 nm vertical cavity surface emitting laser [J]. Journal of Luminescence, 2022, 43(05): 736–744.
46. Liu Songhao. New progress of fiber laser [J]. Optoelectronic Technology and Information, 2003, 16(01): 1–9.
47. Snitzer E. Proposed Fiber Cavities for Optical Masers [J]. Journal of Applied Physics, 1961, 32(01):36–310.
48. Chadler P J, Field S J, Hanna D C, et al. Ion-Implanted Nd: YAG Planar Waveguide Laser [J]. Electronics Letters, 1989, 25(15): 985–986.
49. Li X, Zou W, Chen J. 34.3fs Pulse Generation in an Er-doped Fibre Laser at 201 MHz Repetition Rate [J]. Electronics Letters, 2015, 51(04): 351–352.
50. Mao D, He Z, Gao Q, et al. Birefringence-Managed Normal-Dispersion Fiber Laser Delivering Energy-Tunable Chirp-Free Solitons [J]. Ultrafast Science, 2022, 22(4): 1–13.
51. Xu Linhai, Wang Yufei, Jia Yufei, et al. Research progress of low coherence laser [J]. Journal of Optics, 2021, 41(08): 110–123.
52. Tang Xiaodong, Zeng Qingji, Jin Yaohui, et al. Optoelectronic devices in all-optical communication network and their latest progress (1)[J]. Optical Communication Technology, 2001, 25(02): 90–94.
53. Zhao Chenchen. Research and design of high-speed modulation technology for semiconductor lasers [D]. Xi'an: Xi'an University of Technology, 2016.
54. Konnerth K, Lanza C. Delay Between Current Pulse and Light Emission of a Gallium Arsenide Injection Laser [J]. Applied Physics Letters, 1964, 4(07): 120–121.
55. Goldstein B S, Welch J D. Microwave Modulation of a GaAs Injection Laser [J]. Proceedings of the IEEE, 1964, 52(06): 715–718.
56. Morton P A, Logan R A, Tanbun-Ek T, et al. 25 GHz Bandwidth 1.55 μm GaInAsP P-Doped Strained Multiquantum-Well Lasers [J]. Electronics Letters, 1992, 28(23): 2156–2158.
57. Matsui Y, Murai H, Arahira S, et al. 30-GHz Bandwidth 1.55μm Strain-Compensated InGaAlAs-InGaAsP MQW Laser [J]. IEEE Photonics Technology Letters, 1997, 9(01): 25–28.
58. Sasada N, Nakajima T, Sekino Y, et al. Wide-Temperature-Range (25–80°C) 53-Gbaud PAM4 (106Gb/s) Operation of 1.3μm Directly Modulated DFB Lasers for 10km Transmission [J]. Journal of Lightwave Technology, 2019, 37(07): 1686–16810.
59. Liu G, Zhao G, Sun J, et al. Experimental Demonstration of DFB Lasers with Active Distributed Reflector [J]. Optics Express, 2018, 26(23): 29784–29785.
60. Matsui Y, Schatz R, Pham T, et al. 55 GHz Bandwidth Distributed Reflector Laser [J]. Journal of Lightwave Technology, 2017, 35(03): 397–403.
61. Yamaoka S, Diamantopoulos N P, Nishi H, et al. Directly Modulated Membrane Lasers with 108 GHz Bandwidth on A High-Thermal-Conductivity Silicon Carbide Substrate [J]. Nature Photonics, 2020, 15(01): 28–35.
62. Tian Qi, Han Yu, Zhang Ruigang, et al. Research on high-speed direct modulation semiconductor laser [J]. Optical Communication Research, 2022, 5(8): 1–16.
63. Heidari E, Dalir H, Ahmed M, et al. Hexagonal Transverse Coupled Cavity VCSEL Redefining the High-Speed Lasers [J]. Nanophotonics, 2020, 9(11): 4743–4749.

64. Sun Changzheng, Yang Shuhan, Xiong Bing, et al. Research progress of high-speed electro-absorption modulation laser [J]. China Laser, 2020, 47(7): 30–39.
65. Zhao H W, Pinna S, Sang F Q, et al. High-Power Indium Phosphide Photonic Integrated Circuits [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2019, 25(6): 4500411.1–4500411.10
66. Liang S, Lu D, Zhao L J, et al. Fabrication of InP-Based Monolithically Integrated Laser Transmitters [J]. Science China Information Science, 2018, 61(8): 080405.1–080405.110.
67. Li Sunjun, Hao Yongxin, Zhang Ling, et al. Overview of lithium niobate crystal and its application [J]. Journal of Synthetic Crystals, 2020, 49(06): 947–964.
68. Wang Weidong, Wu Boyu, Li Dejie, et al. Ti:LiNbO3 M-Z intensity modulator for long-distance high-speed optical fiber communication system [J]. Journal of Communication, 1994, 15(06): 52–58.
69. Zhang Bing, Wu Boyu, Zhou Weiqin, et al. 40GHz lithium niobate electro-optical modulator [J]. Optoelectronics Laser, 2001,12 (11): 1199–1201.
70. Gao Zhihui, Gan Qingyun, Wan Junkang. Optimal design of a novel ultra-wideband LiNbO3 electro-optic modulator [J]. Journal of Photonics, 2005, 34(03): 365–369.
71. Yang F, Fang X, Chen X, et al. Monolithic Thin Film Lithium Niobate Electro-Optic Modulator with over 110 GHz Bandwidth [J]. Chinese Optics Letters, 2022, 20(02): 022502.1–022502.5.
72. Wang C, Zhang M, Chen X, et al. Integrated Lithium Niobate Electro-Optic Modulators Operating at CMOS-Compatible Voltages [J]. Nature, 2018, 562(7725): 101–104.
73. He M, Xu M, Ren Y, et al. High-Performance Hybrid Silicon and Lithium Niobate Mach-Zehnder Modulators for 100 Gbit s-1 and Beyond [J]. Nature Photonics, 2019, 13(05): 359–364.
74. Kharel P, Reimer C, Luke K, et al. Breaking Voltage-Bandwidth Limits in Integrated Lithium Niobate Modulators Using Micro-Structured Electrodes [J]. Optica, 2021, 8(03):357–363.
75. Xu M, Zhu Y, Pittalà F, et al. Dual-Polarization Thin-Film Lithium Niobate In-Phase Quadrature Modulators for Terabit-Per-Second Transmission [J]. Optica, 2022, 9(1): 61–62.
76. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress of adaptive optics technology in Optical Wireless Communication system of Xi'an University of Technology [J]. high power laser and particle beams, 2021, 33(08): 30–52.
77. Chen Dan. Research on wireless optical subcarrier modulation and atmospheric influence suppression technology [D]. Xi'an: Xi'an University of Technology, 2012.
78. Ke Xizheng. Principle and Application of Wireless Optical Orthogonal Frequency Division Multiplexing [M]. Beijing: Science Press. 2018.
79. Qiu Kun, Liu Xuming, Mei Kejun. Study on High Frequency Response Characteristics of Semiconductor Laser [J]. Journal of University of Electronic Science and Technology of China, 1993, 22(04): 395–3910.
80. Xiao Yong. Simulation and parameter extraction of equivalent circuit model of semiconductor laser [D]. Chengdu: University of Electronic Science and Technology of China, 2012.
81. Duan Hui. Study on response characteristics of semiconductor laser based on rate equation [D]. Qinhuangdao, Hebei: Yanshan University, 2011.
82. Zhang Lu. Research on nonlinear characteristics of laser in wireless laser communication [D]. Xi'an: Xi'an University of Technology, 2016
83. Shi Weijian. Analysis and research of non-limiting QPSK sinusoidal modulation in atmospheric laser communication [D]. Xi'an: Xi'an University of Technology, 2016.
84. Fan Guidong. Research on power control system of semiconductor laser [D]. Xi'an: Xi'an University of Technology, 2016.
85. Sun Jian, Mou Haiwei, Liu Shiqing, et al. Research on half-wave voltage measurement method in electro-optic modulation [J]. College Physics, 2008, 27(10): 40–43.
86. Chen Haibo, She Weilong. Temperature characteristics and optimal design of electro-optic modulator [J]. Journal of Optics, 2004, 24(10): 1353–1358.
87. Yang Xiao. Research on Electro-optic Modulation Technology and Driving Source [D]. Xi'an: Xi'an University of Technology, 2016.

Research Progress on Passive Modulation in Free-Space Optical Communication

5

Contents

Introduction 130
Passive Modulation and Its Performance Index 131
Performance Analysis 131
Modulation-Form Analysis 134
Foreign Research Progress 134
Domestic Research Progress 137
Working Principle of an MRR FSO System 140
MRR FSO-System Composition 140
Main-Parameter Analysis 141
Modulated Retro-Reflector 144
Angular-Reflector MRR 144
Cat's-Eye MRR 145
MRR FSO Application Prospects and Development Trends 150
Applications 150
Development Tendencies 152
Summary 153
References 153

Abstract

Traditional free-space optical communication requires both ends of the communication to be installed at the same time laser transmitting system, laser receiving system and acquisition, tracking and aiming system, resulting in the weight, volume, power consumption, equipment complexity, and technical difficulty of the system have been improved, thus limiting the development of free-space optical communication. Using the passive modulation technology constructed by asymmetric space laser communication link, the design of small volume, light weight, and low power consumption of the communication terminal is successfully realized, which provides convenience for the application of small- and medium-sized space laser communication technology.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_5

In this chapter, the research status of passive modulation at home and abroad is reviewed, which is mainly summarized abroad. In addition, the composition of the reverse modulation system and the working principle of the reverse modulator are described. Common Modulated Retro-Reflector such as Angle reflector reverse modulator and cat's-eye reverse modulator are analyzed. The selection of wavelength and the size of optical receiver aperture are discussed. This chapter summarizes the different implementation types and characteristics of passive modulation, summarizes the application of passive modulation in optical wireless communication, and looks forward to the future development direction and prospect of this technology. In the future, passive modulation can be developed along the direction of high speed, focal plane pixelation, two-way communication, and coherent detection, and there is a lot of room for development. This chapter can provide a reference direction for the research of passive modulation and a new idea for the research of optical wireless communication.

Keywords

Free-space optical communication · Passive modulation · Modulated retro-reflector · Cat's-eye effect

Introduction

Free-space optical communication (FSO) requires a laser-transmitting or laser-receiving system and an acquisition, tracking, and pointing (ATP) system to be installed at both ends of the communication system. This increases the system's weight, volume, power consumption, and technical complexity and is the main reason limiting the application of free-space optical communication.

Passive-modulation technology uses a modulated retro-reflector (MRR) to modulate the incident light and reflect it back to the transmitting end. Thus, an active light source is not required at the data-sending end [1]. MRR FSO technology does not need to install laser-transmitter and ATP systems at both terminals of the communication link; thus, it can greatly reduce the weight, volume, and power consumption of one end of the link.

In addition, a passive-modulated FSO system does not place high requirements on the interrogator ATP system, and a closed link can be formed by placing the interrogator in the MRR field of view. Therefore, the FSO limitation is effectively solved and the system application is more flexible.

This chapter introduces the system structure of passive-modulation free-space optical communication, the working principle and performance index of modulated retro-reflectors, and analyzes the main parameters. In addition, common MRRs are introduced, which are the angular reflector MRR and cat's-eye MRR [2]. The foreign and domestic research statuses are summarized, and finally, MRR FSO technology is prospected.

Passive Modulation and Its Performance Index

Passive Modulation

In the communication-system modulation process, if the light source and modulating signal are at the transmitter end of the system, it is called active modulation. If the light source and modulating signal are not at the same end of the system, it is called passive modulation. Passive-modulation technology is also called modulated retro-reflector technology [3]. Figure 5.1 shows the communication diagram of passive modulation in a free-space laser link.

Performance Analysis

Bit-Error Rate Analysis

The formula for calculating the average bit-error rate is shown in (5.1), where P_e is the instantaneous bit-error rate.

$$\langle P_e \rangle = \int_0^\infty f(\gamma) P_e \mathrm{d}\gamma \tag{5.1}$$

Because on-off keying (OOK) is easier to realize than other modulation methods, it is generally used in MRR FSO systems at this stage [5]. In OOK modulation, "1" means a light is output, "0" means no light is output, and a threshold-detection method is used to detect and demodulate the threshold voltage m at the receiving end. Therefore, the choice of m is closely related to the accuracy of the decision. The bit-error rate obtained by different selected m is also different [6].

When the sent code element is "1," the probability of erroneously receiving "0" (that is, the receiving voltage $V \leq m$) is as follows:

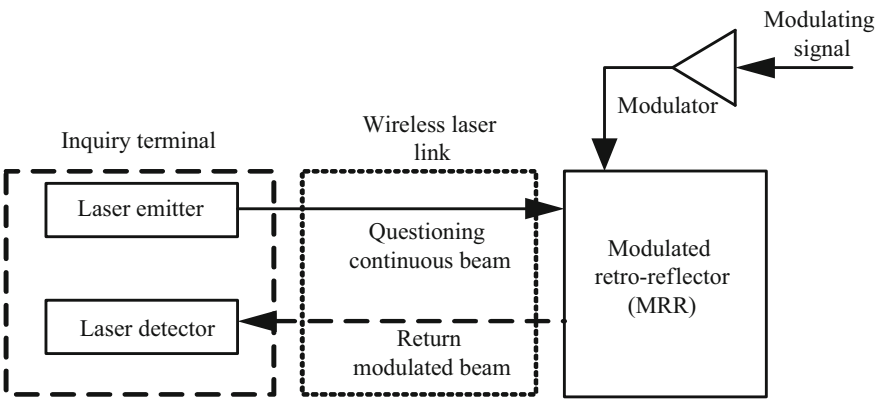


Fig. 5.1 Passive-modulated space laser communication [4]

$$\begin{aligned} P_{e1} &= P(V \leqslant m) = \int_0^m f_1(V)\mathrm{d}V = 1 - \int_m^\infty f_1(V)\mathrm{d}V \\ &= 1 - \int_m^\infty \frac{V}{\boldsymbol{\sigma}_n^2} I_0\left(\frac{aV}{\boldsymbol{\sigma}_n^2}\right) e^{-\left(V^2+a^2\right)/2\sigma_n^2}\mathrm{d}V, \end{aligned} \tag{5.2}$$

where $\boldsymbol{\sigma}_n^2$ is the noise variance; a is the received signal amplitude; and $f_1(V)$ is the probability-density function of the received signal.

The integral value in the above equation can be expressed by the Q function as

$$P_{e1} = 1 - Q\left(\sqrt{2\mathrm{SNR}}, m_0\right), \tag{5.3}$$

where $\mathrm{SNR} = a^2/2\boldsymbol{\sigma}_n^2$ is the signal-to-noise ratio and $m_0 = m/\boldsymbol{\sigma}_n$ is the normalized threshold value.

When the sent code element is "0," the probability of erroneously receiving "1" (receiving voltage $V > m$) is as follows:

$$P_{e0} = P(V > m) = \int_m^\infty f_0(V)\mathrm{d}V = \int_m^\infty e^{-\frac{V^2}{2\sigma_n^2}}\mathrm{d}V = e^{-\frac{m^2}{2\sigma_n^2}} = e^{-\frac{{m_0}^2}{2}}. \tag{5.4}$$

The probabilities of sending code elements "0" and "1" at the modulation end are set to $P(0)$ and $P(1)$, respectively; then, the total bit-error rate of the system is

$$\begin{aligned} P_E &= P(1)P_{e1} + P(0)P_{e0} \\ &= P(1)\left(1 - Q\left(\sqrt{2\mathrm{SNR} = a^2/2\boldsymbol{\sigma}_n^2}, m_0\right)\right) + P(0)e^{-m_0^2/2}. \end{aligned} \tag{5.5}$$

In general, the value is in the range between $\sqrt{2}$ and $\sqrt{\mathrm{SNR}/2}$. When the MRR FSO system works in a high SNR environment, the threshold value is $\sqrt{\mathrm{SNR}/2}$. When $P(0) = P(1) = 0.5$, the bit-error rate can be expressed as

$$P_E = \frac{1}{2}\left[\left(1 - Q\left(\sqrt{2\mathrm{SNR}}, m_0\right)\right) + e^{-\frac{m_0^2}{2}}\right] = \frac{1}{4}\mathrm{erfc}\left(\frac{\sqrt{\mathrm{SNR}}}{2}\right) + \frac{1}{2}e^{-\frac{\mathrm{SNR}}{4}}, \tag{5.6}$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function.

Interruption-Probability Analysis

The interruption probability refers to the probability that the SNR of the received signal from the communication system is lower than the specified threshold γ_{th}. The formula of the interruption-probability pair is shown in Eq. (5.7) [7]:

$$P_{\mathrm{out}} = \Pr\{\gamma \leq \gamma_{\mathrm{th}}\} = \int_0^{\gamma_{\mathrm{th}}} f(\gamma)\mathrm{d}\gamma, \tag{5.7}$$

where $\gamma = \eta^2 h^2 I^2 / 2\sigma_n^2$ is the instantaneous SNR of the received signal; $\overline{\gamma} = \eta^2 \mathrm{E}[h]^2 I^2 / 2\sigma_n^2$ is the average SNR; h is the degraded MRR communication link; η is the photoelectric-conversion coefficient of the photodetector; I is the signal-light intensity when the transmitted item is "1"; and σ_n^2 is the variance.

Channel-Capacity Analysis

The average channel capacity is the maximum error-free information-transmission rate of a communication-system channel. It represents the maximum communication rate that can be transmitted by a communication channel, as shown in Eq. (5.8) [7]:

$$\langle C \rangle = \int_0^\infty B \log_2(1+\gamma) f(\gamma) \mathrm{d}\gamma = \frac{1}{\ln 2} \int_0^\infty B \ln(1+\gamma) f(\gamma) \mathrm{d}\gamma, \tag{5.8}$$

where B is the system bandwidth.

Power Analysis

The laser power of the modulated retro-reflector system [8] and the optical-signal power ultimately received by the active terminal [9] can also be used to measure the communication performance of an MRR FSO system. For a conventional FSO communication system, the attenuation of the system laser power will decrease with $e^{-\alpha R} R^{-2}$. However, for a passive-modulated optical wireless communication system, because its transmission is in accordance with the "back to the original" method, its transmission path is a two-way loop. Therefore, the attenuation of its laser power decreases with $e^{-2\alpha R} R^{-4}$ [10].

The received power expression is shown in Eq. 5.9:

$$P_\gamma = P_t \frac{D_{\mathrm{ret}} D_{\mathrm{rec}} T \beta \gamma}{\Omega_t \Omega_{\mathrm{ret}} R^4}, \tag{5.9}$$

where P_γ is the power of the modulated signal, P_t is the laser power of the transmitting terminal, D_{ret} is the caliber of the MRR, D_{rec} is the caliber of the receiving optical system, $T = \mathrm{e}^{-\alpha R}$ is the air-attenuation constant, β is the modulation coefficient of MRR, γ is the scintillation coefficient of the laser beam after passing through turbulent atmosphere, Ω_t is the unmodulated laser-divergence angle, and Ω_{ret} is the modulated laser-divergence angle.

According to the MRR FSO link-communication process, the optical-signal power finally received by the active terminal is shown in Eq. (5.10):

$$P_{\mathrm{rec}} = \frac{P_t \tau_t \tau_{\mathrm{atm}}^2 D_{\mathrm{rec}}^2 \tau_{\mathrm{rec}} \tau_{\mathrm{MRR}}^2 \rho_{\mathrm{MOD}} D_{\mathrm{ret}}^2 (C_{\mathrm{MOD}} - 1)}{R^4 \Omega_t^2 \Omega_{\mathrm{ret}}^2 C_{\mathrm{MOD}}}, \tag{5.10}$$

where P_t is the laser power at the transmitting end; τ_t is the transmittance of the transmitting optical system; τ_{atm} is the atmospheric transmissivity; τ_{rec} is the transmittance of the receiving optical system; ρ_{MOD} is the surface reflectance of the

photoelectric modulator; τ_{MRR} is the MRR optical-system transmittance; and C_{MOD} is the modulator modulation contrast.

Modulation-Form Analysis

Four types of modulation are used in optical wireless communication systems. They are on-off keying (OOK), pulse-position modulation (PPM), phase-shift keying (PSK), and frequency-shift keying (FSK) [11].

OOK modulation is the most commonly used method; however, in an MRR optical wireless communication system, if both parties use OOK modulation, the communication rate will be greatly restricted and the bit-error rate will increase.

The principle is that the active end of the system sends the "0" code. When the MRR end of the system needs to send the"1" code, an error code will appear because the MRR end omits the laser; hence, the bit-error rate is increased by a quarter.

If the transmission rate is reduced by sending the signal pulse "1" using all-optical storage, the communication rate of the uplink will be reduced to $1/N$ of the communication rate of the downlink. The value of N is determined according to the maximum number of concatenated "0"s when coding. Consequently, it is very restrictive when both sides of the communication adopt OOK modulation in the MRR form.

Therefore, in recent years, many domestic and foreign scientific-research institutions and researchers have been committed to the study of duplex-modulation technology. Using duplex-modulation technology in MRR FSO communication systems can further improve the system transmission rate, increase the spectrum utilization, and reduce the waste of communication resources.

The duplex-modulation technology in MRR FSO communication systems can be divided into two types. One separates the light source and MRR information according to the demodulation orthogonality; it includes PPM [12], carrier modulation [13], and polarization modulation [14]. The other type allocates specific communication resources to the uplink and downlink through separate multiplexing technology to correctly extract the MRR terminal information. It includes time-division multiplexing [15], electrical-domain frequency-division multiplexing, and optical wavelength-division multiplexing [16].

Foreign Research Progress

In 1994, Swenson et al. [17] proposed a model of an MRR FSO system. To verify the feasibility of this communication model, Utah State University and Phillips Air Force Laboratory conducted a communication experiment over a distance of 4 km.

In 1996, the Aerospace Dynamics Laboratory of Utah State University and Swenson carried out an MRR FSO communication experiment between a high altitude and the ground [18]. In this experiment, a ferroelectric liquid-crystal angular

reflector was used to transmit data over a communication distance of about 30 km with a communication rate of 20 Kbit/s.

In 1998, Gilbreath et al. from the United States Naval Research Laboratory experimentally verified an MRR system based on multiple quantum wells (MQWs) [19]. The experiment transmitted image data over 17 m, with a communication rate of 460 Kbit/s. The experiment verified that the MQS-based MRR FSO communication system could transmit data.

In the fall of 1999 and the winter of 2000, Gilbreath et al. experimentally verified MRR FSO communication between rotorcraft unmanned aerial vehicles (UAVs) and the ground in Chesapeake Bay, USA [20]. The experiment transmitted video data over a communication distance of 35–65 m, with a vertical height of 35 m, and a communication rate of 910 Kbit/s.

In 2002, Mahon et al. [21] proposed a new modulation scheme using a cat's-eye MRR. The scheme could simultaneously modulate and reflect incident light. According to this method, the United States Naval Research Laboratory developed an MRR based on the cat's-eye structure and used this MRR in the Chesapeake Bay MRR FSO field experiment. The data transmission had a communication distance of 1 km and a communication rate of 10 Mbit/s.

In 2003, Rabinovich et al. [22] studied a new type of multi-quantum-well modulated retro-reflector. The MQW MRR used a cat's-eye reverse reflector, and the MQW was located on the cat's-eye focal plane. In this study, the size of the modulator was decoupled from the size of the optical aperture to produce a higher data-transmission rate of 50 Mbit/s.

In 2003, Gilbreath et al. [23] carried out an MRR FSO outfield communication experiment in Chesapeake Bay. In this experiment, the data-transmission rate between a UAV and the ground reached 5 Mbit/s. A good bit-error rate was achieved within a communication distance of 30–100 m.

In 2004, Shay et al. [24] from the University of New Mexico used an MRR composed of liquid-crystal materials and angular reflectors to construct an MRR FSO communication system that could realize full-duplex communication with a communication rate of 20 Kbit/s.

In 2006, Plett et al. [25] used an MRR array composed of three MRRs to carry out an MRR FSO outfield-communication experiment in Chesapeake Bay. In this experiment, the communication distance between the two coasts was about 16 km and the communication rate was 2 Mbit/s.

In 2007, Öhgren et al. [26] designed an MRR FSO system using an MQW modulator and a large field-of-view cat's-eye optical system, in which the bandwidth of the MQW modulator exceeded 10 MHz. Multilevel modulation in the MRR FSO was studied, and 90-m indoor and 160-m outdoor transmission experiments were conducted.

In 2008, Rabinovich et al. [27] conducted a sea-surface communication experiment between ships in the Pacific Ocean. The 1550-nm band was used between two moving ships to verify the reliability of MRR FSO communication on the sea surface, when both ends of the communication were in a non-stationary state. An

MRR FSO communication experiment with a communication distance of 4.5 km and communication rate of 2 Mbit/s was realized.

In 2009, Gilbreath et al. [28] conducted an MRR FSO outfield communication experiment on the ground. The experiment transmitted data over a maximum communication distance of 2.5 km at a communication rate of 2 Mbit/s.

In 2010, Rabinovich et al. [29] conducted a communication experiment on the scintillation index in an MRR FSO communication link in China Lake, California. The scintillation index of the MRR FSO communication link was compared with that of a traditional FSO communication link. The results showed that the MRR FSO communication system was more susceptible to atmospheric turbulence.

In 2012, Goetz et al. [30] studied the performance of a micro-UAV modulated retro-reflector laser communication system. In this research, one-to-one and one-to-many communications were carried out using the modulated retro-reflector. The results showed that the application of MRR FSO technology to UAVs could greatly improve the transmission efficiency.

In 2013, Rosenkrantz et al. [31] studied an innovative modulated reflector for free-space optical communication. In this study, an innovative device consisting of a ferroelectric film embedded with nanoparticles was used to construct an MRR, and an asymmetric communication link was established between a base station and multiple robots or astronauts. The research showed that in the entire communication-link range, the bit-error rate was the best when the transmission power was 65 mW.

In 2014, Thueux et al. [32] of Sweden and AIR BUS of France jointly carried out an MRR field-communication experiment between an unmanned aerial vehicle (UAV) and the ground. This experiment transmitted data over a communication distance of 1.2 km at a communication rate of 20 Mbit/s using an MRR application in the UAVs.

In 2016, Anonymous researchers [33] studied a new space optical-communication technology, including communication from earth to a spacecraft in earth orbit and in deep space, such as the moon and Mars. In this study, a new laser-beam pointing technology was used to improve the optical-data transmission of spacecraft. It amplified the laser power received by the modulating reflector on a low-orbit satellite and improved the pointing process between the earth and the spacecraft terminal.

In 2018, Born et al. [34] studied all-optical retro-modulation (AORM) for free-space optical communication and proposed a device and system architecture for free-space optical and optical wireless communication at high data rates over multiway links. The AORM architecture was manufactured and tested and ultimately proved able to meet a total data-rate transmission at the Tbit/s level to achieve multi-directional free-space optical communication.

In 2019, Korotkova et al. [35] from the University of Miami conducted an experimental study on the polarization characteristics of the received signal of an MRR FSO communication system. This study demonstrated that the polarization characteristics of return waves in a single-station dual-pass optical system with a

glass angular-cube mirror were modulated in a complex manner in a clear-air turbulent atmosphere.

In 2020, El Saghir et al. [36] conducted a unified performance analysis of MRR communication systems in all turbulent states, modeled by a generalized Malaga distribution for free-space optical communication links. The probability-density function, cumulative-distribution function, moment-generation function, closing expression of the system fading channel, and system receiving SNR were obtained, and the asymptotic expressions of the average bit-error rate and interruption probability were studied. The results showed that the asymptotic expressions were consistent with those of weak and moderate turbulence at high SNR. With the decrease of the influence of atmospheric turbulence, the performance of the average bit-error rate was improved and the interruption probability was decreased.

In 2021, El Mashade et al. [37] studied the effect of a non-zero axis pointing error on the performance of an MRR FSO communication system. In this study, the mean bit-error rate, system interruption probability, and asymptotic performance of the system's ergodic capacity at low and high SNR were derived, by considering the non-zero displacement of the beam from the center of the optical receiver. The results showed that the performance of the MRR FSO system decreased exponentially with the increase of axial inclination.

In 2021, Trinh et al. [38] conducted a statistical experiment on the experimental channel of a UAV ground-reflecting FSO link and fine-tracking system. In this experiment, a UAV equipped with a modulation mirror was used to modulate the incident beam, and the modulated beam was reflected back along the same path, so that an FSO communication link with the UAV could be established. This technology provided a practical reference for the theoretical performance analysis and engineering design of a UAV reflection FSO system.

Domestic Research Progress

In 2007, Zhang [39] studied optical reverse modulators based on giant magnetostrictive material (GMM) and silicone gels. In this study, the coupling of GMM and silicone gel was used as a reflecting surface on the angular reflector, and the performance of the diagonal reflector was controlled by adjusting the magnetization state of the giant magnetostrictive rod to generate a relatively small displacement. The results showed that under a voltage drive of 1–15 V, the displacement was 0.1–1 μm. In this case, the displacement variation met the modulation condition, which verified the feasibility of the MRR-system design scheme using GMM.

In 2010, Ke et al. [40] from Xi'an University of Technology studied the effect of light beams on MRR FSO link performance and established a mathematical model of the outgoing light beam after a Gaussian light beam was reflected by the MRR of a corner-cube retroreflector (CCR). The results showed that the laser beam reflected by the CCR had a certain channel-compensation effect on the turbulent channel, and the reflected multi-beam could effectively reduce the received SNR instability.

In 2013, Shan et al. [41] studied the effect of the defocus of a cat's-eye optical system on the terminal scintillation in an MRR FSO system. The results showed that a certain amount of defocusing was needed to minimize the scintillation index (SI) with the changing turbulence and transmission distance. As the defocus increased, the SI tended to first decrease and then increase; the larger the distance and turbulence, the larger the SI.

In 2014, Yang et al. [42] designed a cat's-eye MRR laser-communication receiving-amplifier circuit. Its bandwidth could reach 140 MHz, the signal gain was greater than 70 dB, and the total noise current was about 1.34×10^{-7} A. The results showed that the receiving-amplifier circuit met the conditions of low noise, high gain, and high bandwidth in cat's-eye MRR laser communication.

In 2015, Qiu et al. [43] proposed a single light-source full-duplex MRR FSO scheme, which adopted OOK modulation at the data-transmission end and amplitude modulation at the MRR end. After the experiment, the information-communication process of the scheme was analyzed, and a short-distance MRR single light-source full-duplex optical communication experiment was completed in an indoor environment.

In 2015, Xu et al. [44] studied an MRR optical communication system based on acousto-optic modulation. The experimental results showed that this communication system achieved a communication rate of 118.2 Kbit/s and high carrier frequency of 1 MHz, and the bit-error rate was much lower than 10^{-6} for the data transmission.

In 2016, Ke et al. [45] from Xi'an University of Technology established an equivalent model of the defocusing structure of the cat's-eye target using geometric optics and studied the factors affecting the characteristics of reverse-modulated reflected light. The results showed that the MRR receiving distance and internal-structure parameters had obvious effects on the power density and reflected-spot radius of the modulation system. This effect changed regularly. Reducing the lens focal length, increasing the lens aperture, and increasing the receiving distance all accelerated the change of reflected power with the defocus quantity.

In 2017, Ke et al. [46] from Xi'an University of Technology designed an indoor full-duplex communication system with a single light source, which was based on a defocusing cat's-eye MRR scheme. When the communication distance was 3.0 m, the upstream data transmission was 2.0 Kbit/s and the downstream data transmission was 5.0 Mbit/s.

In 2018, Zhang [47] designed a visible-light full-double-mode communication system and proposed a new visible-light reverse-transmission scheme, based on the defocusing cat's-eye. Using an upstream modulator based on the defocusing cat's-eye MRR principle and a signal-processing circuit to remodulate the downlink optical signal, the experiment realized a visible-light bidirectional communication link with a single-LED light source.

In 2019, Chen et al. [48] studied data transmissions in MRR laser-communication links and presented a transmission model based on combinatorial optimization theory. Through research on this topic, the relationship between the data-transmission characteristics and various factors in the MRR optical-fiber

communication system was found. The model effectively improved the success rate of the data transmission in the system and improved the system data-transmission rate.

In 2019, Pan et al. [49] studied the atmospheric-turbulence effect and its control method in an MRR optical-communication echo-reflection link. Based on the phase-screen method, this project constructed traditional FSO and MRR FSO transmission models of Gaussian beams under weak turbulence conditions and theoretically analyzed the flicker characteristics of the Gaussian beam during MRR FSO transmission. The results showed that atmospheric turbulence had a great influence on the MRR optical-communication echo-reflection link, and the influence was much greater than that on the one-way link in traditional FSO.

In 2020, Duan et al. [50] calculated the energy of the cat's-eye laser echo after attenuation in the atmosphere and analyzed the transmission distance, wavelength of the emitted laser, angle of the incident laser, cat's-eye aperture, and other factors and determined the influence of the above factors on the signal-to-noise ratio at the receiving end. The results showed that the larger the transmission distance and the larger the cat's-eye and receiving apertures, the larger the SNR at the receiving end. In contrast, the larger the laser incidence angle, the smaller the signal-to-noise ratio. Moreover, a long-wavelength laser could improve the signal-to-noise ratio of the link.

In 2020, Li [51] proposed an MRR FSO communication system based on an adaptive decision threshold and compared it with an MRR FSO communication system based on an optimal decision threshold. The experiments showed that the MRR FSO communication system with the adaptive decision threshold could communicate with simple equipment and no prior channel information. The difference in the bit-error rate between the two systems was only 1.3 dB.

In 2021, Jia et al. [52] designed an intersatellite laser-communication system for small satellite formations based on a reverse modulator. In this scheme, by constructing a complete satellite navigation-system model, the influence of random pointing errors on the average bit-error rate and average interrupt rate of an MRR communication system was studied. The experimental results showed that the average interrupt rate and the average bit-error rate could be better than 10^{-6} with a 1-Gbit/s communication rate and a 200-km communication distance as experimental conditions.

In 2021, Fang et al. [53] designed an MRR optical-communication signal determination theory based on the K-means algorithm. Considering the characteristics of the received sample points, the sample points in the window were dynamically and adaptively classified according to the similarity of their characteristics, which greatly improved the signal-decision accuracy. The results showed that this signal-decision method based on the K-means algorithm had better error performance in turbulent channels.

In 2022, Chen et al. [54] designed a combinational-logic reverse-pulse bit-modulation scheme to eliminate the downlink-signal interference to the uplink signal, which reduced the bit-error rate of a single-light MRR FSO communication

system. The results showed that the gain of this system was about 5 dB. The scheme obtained better bit-error rate performance when the optical transmitter, optical receiver, and MRR used the reverse-pulse bit-modulation scheme.

In 2022, Xie et al. [55] used the phase-screen method to study the effect of atmospheric turbulence on an MRR optical-communication system. Based on the model, the scintillation index (SI) was calculated and the aperture average effect of the MRR terminal was analyzed. The SI values of the MRR link and direct link were compared in a 506-m transmission link. The experimental results showed that the experimental data were in good agreement with the numerical simulation results, so the feasibility of the theoretical model was verified.

Working Principle of an MRR FSO System

MRR FSO-System Composition

A modulated retro-reflector system can include MRR FSO communication systems based on liquid crystal, piezoelectric ceramics and micro-electromechanical structures, acousto-optic modulators, multi-quantum wells, and advanced modulation formats [51]. From the perspective of a free laser-communication link, an MRR FSO system includes two parts: the drive end and the cat's-eye reverse end.

The drive end of the MRR FSO is composed of an optical base and servo turntable. The optical base consists of a communication transmitting and receiving unit, beacon transmitting and receiving unit, and auxiliary-target viewing unit. The optical base integrates each optical path and ensures their coaxiality. The communication receiving unit receives the communication light reflected back by the reverse end and provides information for tracking the servo turntable. The servo turntable mainly receives the position information given by the beacon and ensures the stability of the communication link.

The reverse end of the MRR FSO consists of three parts: a cat's-eye optical system, laser modulator, and signal modulator. Among them, the cat's-eye optical lens adopts an improved spherical design. The cat's-eye optical lens can be used to receive large-field optical signals; then, the incident light is transmitted to the drive end, according to the original path through the lens. The signal modulator modulates and encodes the information in the link, and the control modulator modulates the incident laser.

The beacon light emitted by the drive end of the MRR FSO returns along the original path of incidence after passing through the cat's-eye optical system. The position information of the reverse end is detected by the beacon receiving unit in the optical base, and the servo turntable accurately locates the reverse end. The cat's-eye optical lens gathers the communication light emitted by the drive end onto the modulator target surface, and the modulator loads the information onto the communication light for transmission by changing the intensity of the incident light.

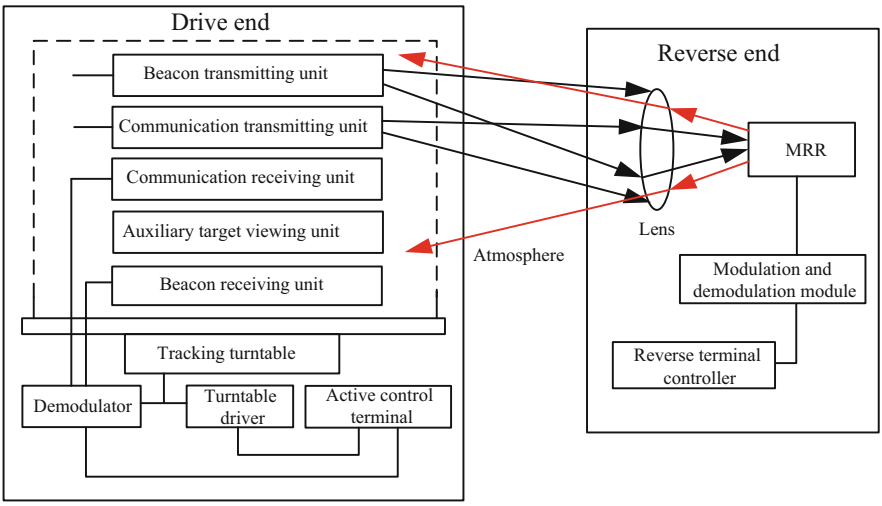


Fig. 5.2 Composition of the modulated retro-reflector laser communication system [56]

It returns the modulated communication light to the receiving unit of the drive end through the original path of the cat's-eye lens, and the drive end demodulates it to obtain the information transmitted by the reverse end, completing the modulated retro-reflector [56]. The composition of the MRR FSO system is shown in Fig. 5.2.

Main-Parameter Analysis

Compared with a traditional FSO system with the same laser-communication link distance, the communication laser needs to pass through the atmospheric channel twice, in accordance with the "back to the original" method, and the transmission loss is squared. To ensure that the drive end of the MRR FSO accurately tracks the reverse end of the non-tracking unit, the following parameters are analyzed.

Wavelength Choice

The system operating wavelength has a very important impact on the communication performance. Under the same intensity turbulence, the longer the wavelength used, the less the impact will be. Limited by the absorption window of the atmosphere and considering the technical complexity of the laser [34], two laser bands are commonly used in laser communication, 800 nm and 1550 nm.

The scattering effect of the atmosphere in these two bands is manifested in the form of Mie scattering, which occurs when the diameter of the particles in the atmosphere is equivalent to the wavelength of the radiation. Mie scattering is related to the atmospheric visibility, wavelength, and correction factor. In general, it is represented by the attenuation coefficient, and its expression is

$$\sigma = \frac{3.912}{V_m}\left(\frac{\lambda}{550}\right)^{-q}, \tag{5.11}$$

where V_m is the atmospheric visibility; λ is the wavelength; and q is the correction factor. The relationship between the size of the correction factor and atmospheric visibility is

$$q = \begin{cases} 1.6 & (V_m > 50\,\text{km}) \\ 1.3 & (6\,\text{km} < V_m \le 50\,\text{km}) \\ 0.16V_m + 0.34 & (1\,\text{km} < V_m \le 6\,\text{km}) \\ V_m - 0.5 & (0.5\,\text{km} < V_m \le 1\,\text{km}) \\ 0 & (V_m \le 0.5\,\text{km}) \end{cases}. \tag{5.12}$$

In actual communication, the relationship between the laser-power attenuation and the change of distance is shown in Eq. (5.13).

$$\beta = -10\sigma \cdot L \tag{5.13}$$

According to the above formula, Fig. 5.3 shows the relationship diagram of the laser-power attenuation with distance when the wavelength is 800 nm (solid line) and 1550 nm (dashed line). As shown in the figure, the laser in the 1550-nm band decreases 0.7-dB per kilometer less than the laser in the 800-nm band; hence, the 1550-nm band is typically used in communication.

Beam-Divergence Angle

The divergence angle of the laser beam can represent the velocity of the beam from the waist. To meet the communication and application requirements, a low beam-

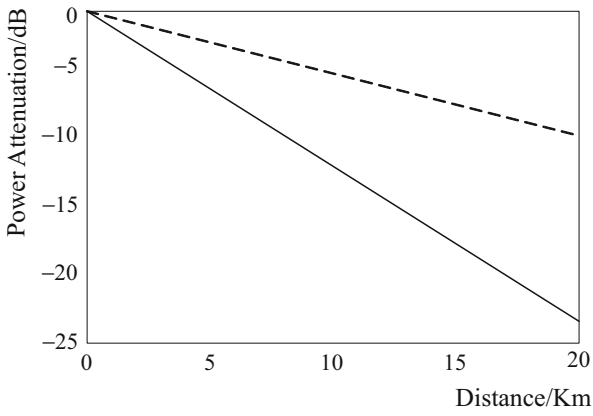


Fig. 5.3 Laser-power attenuation with distance [56]

divergence angle is required in FSO communication. The beam-divergence angle is mainly limited and affected by the following three factors:

1. Free-space loss: The free-space loss is the energy loss of an optical signal propagating in atmosphere. For an FSO system, the free-space attenuation is directly proportional to the square of the communication beam-divergence angle.
2. Communication optical-diffraction limit: The diffraction limit is related to the optical wavelength and communication optical-emission aperture. The specific relationship is shown in Eq. (5.14):

$$\theta = \frac{2.44\lambda}{D}, \tag{5.14}$$

where θ is the diffraction limit; λ is the wavelength; and D is the communication optical-emission aperture.

Registration precision: The alignment accuracy is related to the received power attenuation of the communication system, and the received power-loss coefficient can measure the alignment accuracy of the system. The loss coefficient is related to the tracking error and beam-divergence angle. The specific relationship is shown in Eq. (5.15):

$$\alpha = e^{-8\left(\frac{\theta_e}{\theta_T}\right)^2}, \tag{5.15}$$

where α is the received power-loss coefficient; θ_e is the tracking error; and θ_T is the beam-divergence angle. According to Eq. (5.15), when θ_T decreases, α increases, and the laser power loss is greater.

Optical-Receiver Diameter

Because the atmosphere is the communication carrier of an FSO system, the FSO system is greatly affected by weather in the communication process. To improve the communication efficiency and receiving power of the system, when an FSO system is applied in long-distance communication, the receiving power of the system is usually increased by the square in large-aperture receiving.

When communicating with the atmosphere as the carrier, a flickering smooth atmosphere is called the aperture-smoothing effect [57]. It can be represented by the aperture-smoothing factor F, and the specific relationship is shown in Eq. (5.16):

$$F = \frac{16}{\pi D^2}\int_0^D \frac{C_I(\rho)}{C_I(0)}\left\{\arccos\left(\frac{\rho}{D}\right) - \frac{\rho}{D}\left[1 - \left(\frac{\rho}{D}\right)^2\right]^{1/2}\right\}\rho \mathrm{d}\rho, \tag{5.16}$$

where $C_I(\rho)$ is the light-intensity covariance function; $C_I(0)$ is the light-intensity variance function; and D is the receiver aperture.

Modulated Retro-Reflector

The modulated retro-reflector is the main component and core of a modulated retro-reflector system. The optical modulator and the reverse reflection system are the main components of the MRR. The signal is processed in the optical modulator and converted into a specific driving signal. This changes the properties of specific materials or devices, thereby changing the laser properties interacting with them. There are two common MRRs in MRR communication: angular-reflector and cat's-eye [2].

Angular-Reflector MRR

The angular-reflector MRR consists of two parts [1], namely, the angular reflector and optical modulator. The angular reflector consists of three highly reflective planes that are perpendicular to each other. Its geometric characteristics can improve the reverse-reflection performance of the incident beam [47].

The angular-reflector MRR uses two different modulation principles. One changes the intensity of the reflected beam by changing the reverse-reflection characteristics of the angular reflector. The other places a special light modulator in the angular reflector and modulates the intensity of the reflected light signal through the photoelectric characteristics of the modulator.

These two MRR types change the reflection characteristics of the reflecting-surface and transmission-switching MRR systems. The MRR that changes the reflection characteristics of the reflector uses a micro-electromechanical system (MEMS) to change the reflector structure. The transmission-switching MRR modulates the reflected light by controlling the switching optical modulator and utilizing the reverse-reflection characteristic of the angular reflector [58]. Figure 5.4 shows a schematic diagram of the angular reflector.

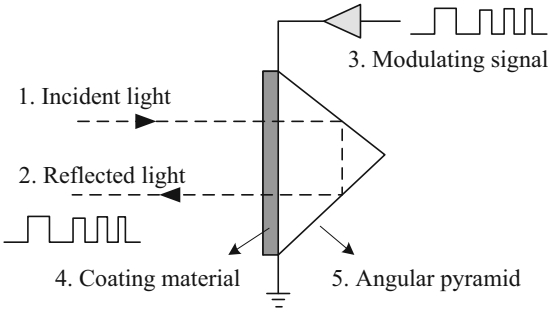


Fig. 5.4 Schematic diagram of an angular-reflection reverse modulator [58]

Cat's-Eye MRR

Cat's-Eye Effect

When a beam of light illuminates a diffused object, its reflected light is reflected in multiple directions, typically with a reflectivity in the range of 0.2–0.5 [59]. When the distance between the light source and the object is large, the reflected light is very weak. However, when the light from the source strikes a photoelectric device with an optical window, the intensity of the reflected light varies greatly.

The position of the photosensitive surface on the photoelectric sensor is near the focal plane of the objective lens of the optical system; therefore, the reflected light is distributed in a small three-dimensional angle range, forming a large reflection cross-section. Just as light from a light source enters a cat's eye, it is focused through the pupil of the cat's eye to the fundus [60]. The cat's-eye effect occurs when an incoming beam of light bounces off the fundus and returns along its original path.

Owing to the cat's-eye effect, the cat's-eye optical system has the same reverse-reflection characteristics as the angular reflector; thus, the cat's-eye system can be combined with an electro-optic modulator to build a cat's-eye MRR.

The cat's-eye MRR consists of two parts: a focusing lens and a modulating reflector. When the laser from the transmitter is transmitted to the cat's-eye lens, the cat's-eye optical system focuses the emitted laser on the reflecting surface of the modulator. The incident light is reflected on the target surface of the modulator of the cat's-eye optical system in accordance with the light characteristics along the original path. The reflected laser passes through modulation and carries useful information back to the light source.

The cat's-eye MRR is passively controlled by the light source. Only when the light source irradiates the optical system of the cat's-eye structure will the incident light be affected by the cat's-eye effect and reflect back to the source in the original manner. It is also known as a passive modulator [3]. Figure 5.5 shows a typical cat's-eye structural physical model.

As the optical antenna of an MRR optical wireless communication system, the performance of the cat's-eye optical structure is directly related to the stability of the

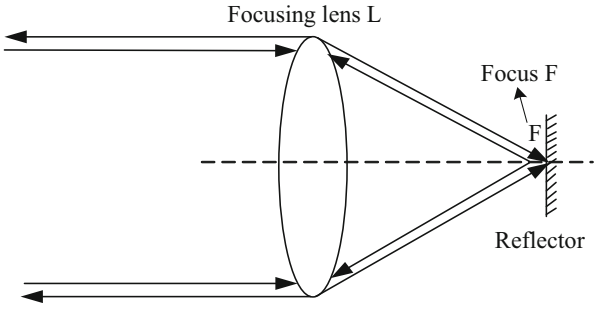


Fig. 5.5 Physical model of a cat's-eye structure [23]

communication link. Therefore, the cat's-eye MRR FSO system in the communication process must meet the following requirements:

1. The cat's-eye optical system must be in accordance with the reverse-reflection characteristics of the system;
2. The cat's-eye optical system must ensure that the divergence angle of the reflected light is as close as possible to the diffraction limit;
3. The F number of the cat's-eye optical system must be as small as possible, where F is the ratio of the equivalent focal length to the optical aperture of the cat's-eye optical system.

Echo-Power Analysis

The cat's-eye MRR modulates the echo power of the incident light to the cat's-eye structure by changing the defocus of the modulator to realize data transmission and communication.

The emitted laser power of the laser-emitting device is P_0, the divergence angle of the emitted laser beam is θ_0, the effective receiving area is A_r, the link distance at which the communication is completed is R, the cat's-eye effective cross-sectional area is A_s, the reflectance is ρ_s, the effective cross-sectional area of small diffuse reflection targets is A_1, the reflectance coefficient is ρ, the two-way atmospheric transmittance is τ_a, the transmittance of the receiving optical system is τ_r, and the divergence angle of the modulated laser beam is θ_s. The cat's-eye-target echo power is shown in Eq. (5.17) [61]:

$$P_s = P_0 \frac{16\tau_a \tau_r A_r A_s \rho_s}{\pi^2 \theta_0^2 \theta_s^2 R^4}. \tag{5.17}$$

The diffuse earth echo power is shown in Eq. (5.18):

$$P_r = P_0 \frac{\tau_a \tau_r A_r \rho}{\pi R^2}. \tag{5.18}$$

From Eqs. (5.17) to (5.18), the ratio of the echo power of the cat's-eye target to the diffuse earth echo power is

$$\frac{P_s}{P_r} = \frac{16 A_s \rho_s}{\pi \theta_0^2 \theta_s^2 R^2 \rho}. \tag{5.19}$$

From Eqs. (5.17–5.19), it can be seen that the ratio of the echo power of the cat's-eye target to the earth echo power becomes increasingly smaller with increasing distance. Because the earth echo power is inversely proportional to the square of the distance, the echo power of the cat's-eye target is inversely proportional to the fourth power of the distance. When the distance reaches a certain value, the ratio between them will be reduced to zero. At this point, cat's-eye MRR communication cannot be used.

Defocusing the Cat's-Eye MRR

When the target lens is in an ideal state, namely, a non-defocused state, there is a high ratio between the intensities of the reflected and incident lights. However, in practical optical-communication systems, owing to manufacturing errors in equipment and components, the photosensitive surface in the optical instrument is not simply in the focal plane; there is usually a certain amount of defocus. Sometimes, to meet specific needs, a certain amount of defocus will be artificially set.

The defocused cat's-eye MRR is an optical reflector placed around the focal plane of the cat's-eye optical system. The intensity of the reverse reflection is modulated by controlling the positions of the reflector and the focal plane [2]. For forward defocusing, where the reflector is positioned in front of the focal point and the optical reflector in the focal plane moves toward the lens, the amount of defocus is negative, also known as negative defocus. The physical model is shown in Fig. 5.6.

For backward defocusing, where the position of the reflector is behind the focal point and the optical reflector on the focal plane moves away from the lens, the defocus is a positive number, also known as positive defocus. The physical model is shown in Fig. 5.7 [56]. In the figure, δ is the defocusing amount, f is the cat's-eye focal length, and F is the focus.

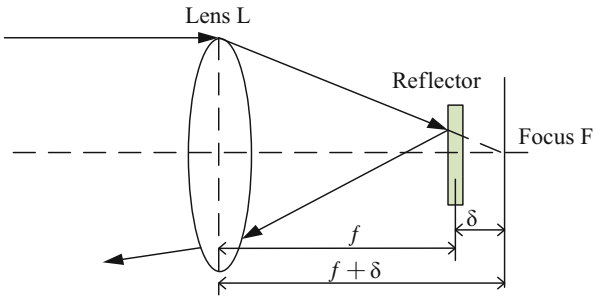


Fig. 5.6 Physical model of forward-defocusing light propagation [56]

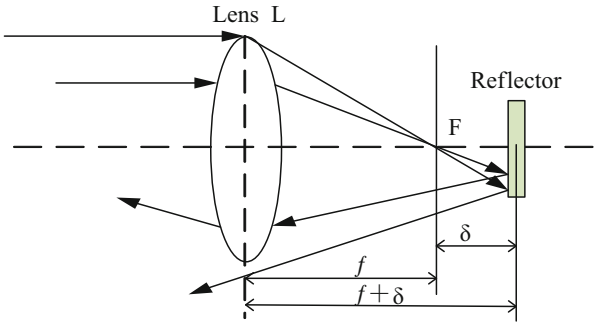


Fig. 5.7 Physical model of backward-defocusing light propagation [56]

Defocusing complicates the reflection characteristics of the laser on the reflecting elements and has some influence on the backward-reflection characteristics and cat's-eye effect [62].

Theoretical Analysis of Non-defocusing

When the cat's-eye optical system is in the non-defocused state, the full field angle of the cat's-eye lens is set as β. Research shows that no beam returns when the incidence angle meets $\theta > 0.5\beta$. When $0 < \theta \leq 0.5\beta$, the light reflects. Therefore, it can be seen that θ must be less than half the field angle, so that the light can reflect backward.

Moreover, when the incident beam is oblique, the divergence angle of the reflected beam will not be affected; however, the change of effective aperture will be affected, and the energy passing rate will be reduced. Within the half-field angle, the effective aperture and the divergence angle of the backward-reflected beam satisfy the following relation:

$$D' = D - 2f\tan\theta \tag{5.20}$$

$$\theta_s = \theta_{in}, \tag{5.21}$$

where D is the cat's-eye-target lens caliber and θ_{in} is the divergence angle of the incident laser beam.

According to Eq. (5.20), when the incident angle is 0°, the incident beams can all return in the same way, which is the cat's-eye effect in the perfect ideal case.

Theoretical Analysis of Defocusing

When the cat's-eye optical system is out of focus, the detector can receive the critical field of view β' of the incident light. When the incident angle satisfies $0 < \theta \leqslant 0.5\beta'$, the beam partially returns. The effective aperture and the divergence angle of the backward-reflection beam satisfies the following relation:

$$D' = (D - 2f\tan\theta)\frac{f+\delta}{f+2\delta} \tag{5.22}$$

$$\theta_s = \theta_{in} + \frac{2\delta}{f^2}(D - 2f\tan\theta)\frac{f+\delta}{f+2\delta}. \tag{5.23}$$

When $\delta = 0$, Eqs. (5.22) and (5.23) can be simplified to Eqs. (5.20) and (5.21), where the forward and backward defocusing satisfy the same formula. The forward defocusing satisfies Eqs. (5.24) and (5.25), and the backward defocusing satisfies Eqs. (5.26) and (5.27) within a range of very small incident angles:

$$D' = D \tag{5.24}$$

$$\theta_s = \theta_{in} + 2\delta D/f^2 \tag{5.25}$$

$$D' = \frac{Df}{f + 2\delta} \tag{5.26}$$

$$\theta_s = \theta_{in} + \frac{2\delta D}{f(f + 2\delta)}. \tag{5.27}$$

From the above formulas, it can be seen that the different focusing points lead to different effects of the two defocusing conditions on the cat's-eye effect. Because the beam propagates according to the diffraction-propagation law and the special propagation law of the laser beam, the far-field characteristics of the reflected beam are more obviously affected by the defocus, and the difference between the two defocus conditions is more obvious.

Electro-Absorption Cat's-Eye MRR

The electro-absorption MRR places an electro-absorption optical modulator at the focal plane of the cat's-eye optical system. Commonly used electro-absorption modulated reflectors include multiple-quantum-well (MQW) modulators and micro-electromechanical system (MEMS) modulators [63]. The MQW and MEMS modulators are shown in Figs. 5.8 and 5.9, respectively.

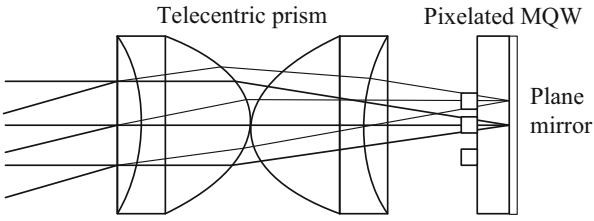


Fig. 5.8 Schematic diagram of multi-quantum-well cat's-eye reverse-modulation optical communication [63]

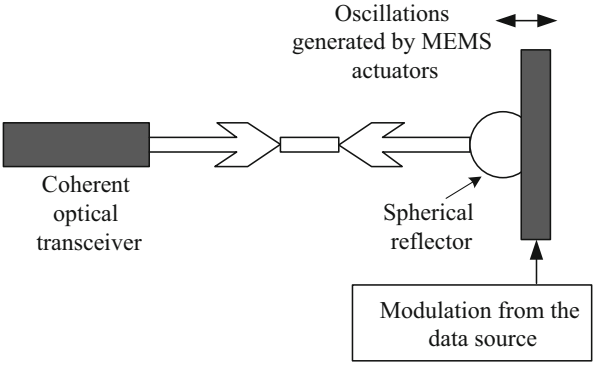


Fig. 5.9 Schematic diagram of microelectromechanical MRR optical communication [63]

The MQW device has the advantages of stability, all-solid-state structure, low driving voltage, low power, and a 40-Gbit/s switching speed. Therefore, in 1998, the US Naval Research Office applied an MQW device to an FSO system and extensively researched its characteristics [64]. The power consumption of the MQW modulator is as shown in Eq. (5.28):

$$P_{\mathrm{MOD}} = \eta C V^2 S, \tag{5.28}$$

where η is the charging circuit, C is the capacitance, V is the driving voltage, and S is the data rate. The charging circuit and communication mode determine the factors. When the charging circuit and communication mode are fixed, the power consumption of an MQW modulator mainly depends on V, followed by C and S.

In addition, the relationship between the maximum modulation rate of the MQW modulator, body resistance, and diameter is shown in Eq. (5.29):

$$S \propto \frac{1}{R_{\mathrm{MOD}} D_{\mathrm{MOD}}^2}, \tag{5.29}$$

where R_{MOD} is the MQW modulator resistor and D_{MOD} is the MQW modulator diameter.

From this formula, the smaller the modulator resistance and diameter, the higher the maximum modulation rate. The modulation rate is an important parameter of the MQW modulator, which directly affects the performance of the inverse modulator. By analyzing the characteristics of the MQW modulator, the key factors of the MQW modulator used in an MRR system are revealed: high switching rate, low power consumption, large field of view, and high modulation efficiency [20].

MRR FSO Application Prospects and Development Trends

Applications

Table 5.1 shows a detailed comparison of the angular-reflector MRR and cat's-eye MRR, according to a study of the system characteristics and communication performance.

MRR FSO communication technology is suitable for high-load satellites, unmanned aerial vehicles, and other platforms. It has been widely used in various fields, such as aerospace communication, maritime-ship communication, atmospheric monitoring, underwater submarine communication, target rendezvous, docking, and identification, and complex electromagnetic-interference environments, making end-to-end communication more reliable. In an MRR optical wireless

Table 5.1 Comparison of two reverse modulators

Angular-reflector MRR	1. Because the angular cone prism is composed of three highly reflective perpendicular planes, it has good reflection characteristics. If the reflectance of one plane is changed, the echo intensity can be changed 2. The attenuated total reflection can be used to fabricate optical MRRs based on angular cone prisms
Cat's-eye MRR	1. The cat's-eye MRR has a strong resistance to electromagnetic interference and can replace a microwave communication link in electromagnetic interference environments 2. Owing to its small size and high capacity, the cat's-eye MRR can monitor the attitude of a target platform and remotely identify the target during rendezvous and docking 3. The directivity of the cat's-eye MRR is very precise, so the communication-safety performance can be guaranteed 4. The transmission rate of a cat's-eye MRR system can reach Gbit/s or even higher, and it can transmit large amounts of multimedia data at the microwave-frequency limit

communication system, this technology plays an important role in the following aspects:

1. Small-platform communication. MRR FSO can minimize the architecture of a terminal on the premise of ensuring the angle of the field of view, and it has the backtrack feature. It has great application prospects in space satellites, space-based unmanned aerial vehicles, underwater vehicles, and other small platforms, as well as communication with ground systems. In addition, by optimizing the terminal architecture, multi-user collaborative communication can be achieved, which is especially suitable for small platforms. This can establish a full-coverage communication network and provide an important foundation for the future development of communication technology.
2. Identity and posture recognition. By transmitting signals outward through MRR, or combining MRRs into an MRR array according to a certain arrangement, remote targets can be identified through a communication link between a micro-MRR terminal and the remote drive end. This method is used to transfer position-feature information during rendezvous and docking. By establishing communication links between MRRs in various characteristic positions, an actively docking aircraft can obtain the attitude information of the docking object and adjust its attitude to dock accurately.
3. Because MRR wireless laser communication has a strong resistance to electromagnetic interference (EMI), MRR FSO can be used to replace the radio-frequency (RF) link in strong EMI environments with explosions and other complex signals. In addition, the transmitting beam-divergence angle is small, difficult to intercept, has good security, and can be used for confidential communications.

Development Tendencies

Because MRR is a rapid-targeting wireless laser-communication technology [65], it can effectively avoid technical problems, such as following and pointing between communication parties. However, traditional MRR technology can only load unilateral information and communicate in one direction. In recent years, many foreign research institutions have been studying the new MRR full-duplex transmission technology (FDX) [66], trying to find a lower power and lightweight FDX transmission scheme.

Based on domestic and foreign achievements in space laser-communication MRR research, it is believed that the future development of space laser communication under MRR will be as follows:

1. Higher speeds. Multi-quantum-well devices are some of the most important devices that can truly realize high-speed MRR spatial optical communication at present. Although it is reported that optical-fiber devices can be used for gigabit data transmissions, this form of MRR is inferior to MRR based on quantum wells because it loses the characteristic of original MRR that required less ATP.

   Furthermore, although the modulation rate of the multi-quantum-well device itself can be tens of gigahertz, the current maximum rate of systems based on multi-quantum-well MRR spatial optical communication systems is only 200 Mbit/s, which leaves extensive room for development.
2. The focal plane can be pixelated. MRRs based on multi-quantum-well devices mainly adopt the cat's-eye structure, which has a relatively small field-of-view angle. Expanding the field-of-view angle has become the focus of attention in various countries. One direct method is to increase the area of multi-quantum-well devices; however, the modulation rate decreases significantly with the increase of area.

   Thus, the single multi-quantum-well device at the focal point of the lens can be expanded into several devices. This forms a pixelated multi-quantum-well focal plane to achieve a large field-of-view angle while keeping the modulation rate constant. However, there is less research in this area. Another future direction of study involves separately controlling each pixel to only work under light irradiation, to reduce the power consumption.
3. Two-way communication. The traditional MRR space optical communication system has unidirectional characteristics; that is, the MRR terminal data are sent to the inquiry terminal. In some cases, two-way communication is required, and traditional MRR technology cannot meet the requirement. Therefore, the MRR terminal is equipped with both the receiving and inquiring ends; thus, two-way data communication can be achieved by a simple modification.

   For example, Changchun University of Science and Technology in China uses dual wavelengths to accomplish two-way MRR data communication. Moreover, the PLA University of Science and Technology suggests using different modulation methods to carry out two-way MRR communication.

4. Coherent detection. Because the beam of an MRR space optical communication system must pass through two optical paths, the system loss multiplies, compared with the traditional space laser communication link, and becomes an important reason to restrict the communication distance. Coherence-detection technology can significantly improve the sensitivity of the interrogator signal and improve the communication distance of MRR spatial optical communication systems.

Summary

This chapter introduced the principle of passive-modulation optical wireless communication technology, explained the structure and working principle of the core part of the MRR passive-modulation system, and analyzed the main parameters. It introduced the angular-reflector MRR and cat's-eye MRR, analyzed the different types of MRR and their principles, and summarized the domestic and foreign research statuses.

In the course of introducing the angular reflector MRR, the MRR system that changed the reflection characteristics of the reflector and the transmission-switch MRR system were introduced. In the course of introducing the cat's-eye MRR, the defocusing cat's-eye MRR and electro-absorption cat's-eye MRR were introduced.

This chapter discussed the main existing problems and possible solutions of MRR optical wireless communication systems. In addition, the MRR development trends in optical wireless communications were discussed. It is believed that MRR FSO will receive increasing attention and play an important role in future communication networks.

References

1. Ke Xizheng, Zhang Huan. Indoor Visible Uplink Communication System Based on Modulated Retro-Reflector [J]. Advances in Laser and Optoelectronics, 2018, 55(02):115–122.
2. Sun Huayan, Zhang Laixian, Zhao Yanzhong, et al. Advances in Modulated Retro-Reflector Free-Space Laser Communication Technology [J]. Advances in Laser and Optoelectronics, 2013, 50(04):33–41.
3. Ke Xizheng, Optical Wireless Communication [M]. Beijing: Science Press, 2016: 54–58.
4. Zeng Zhilong, Hunag Haiyang, Wang Xia, et al. Advances in Modulated Retro-Reflector Space Laser Communication Technology [J]. Optical Communication Technique, 2017, 41(12): 16–110.
5. Li Zhan, Peng Renjun, Wu Jian, et al. Optical Modulated Retro-Reflector Technique [J]. Advances in Laser and Optoelectronics, 2007(11):47–51.
6. Swenson Charles M., Jensen Gary L. Low-power optical transceiver for low earth orbit [J]. Utah State Univ, 1995, 5(07): 2553–2553.
7. Swenson Charles M., Steed Clark A., De La Rue Imelda A., et al. Low-power FLC-based retromodulator communications system [J]. Utah State Univ, 1997, 10(03): 296–311.
8. Gilbreath G C, Rabinovich W S, Mahon R, et al. Modulating retroreflector architecture using multiple quantum wells for free space optical communications [J]. International Conference on Applications of Photonic Technology III, 1998, 7(06): 581–586.

9. Gilbreath G. Charmaine, Rabinovich William S., Meehan Timothy J., et al. Compact lightweight payload for covert datalink using a multiple quantum well modulating retroreflector on a small rotary-wing unmanned airborne vehicle [J]. Naval Research Lab, 2000, 11(10): 4127–4128.
10. Rabinovich W S, Mahon R, Goetz P G, et al. A Cat's eye multiple quantum-well modulating retro-reflector [J]. IEEE Photonics Technology Letters, 2003, 15(03): 461–463.
11. Rabinovich William S., Goetz Peter G., Mahon Rita, et al. Cat's eye quantum well modulating retroreflectors for free-space optical communications [J]. Naval Research Lab, 2003, 11(10): 4975–4975.
12. G. Charmaine, William S. Rabinovich, Timothy J. Meehan, et al. Progress in development of multiple-quantum-well retromodulators for free-space data links [J]. Optical Engineering, 2003, 42(06):332–399.
13. Shay Thomas M., MacCannell John A., Garrett Christopher D., et al. First experimental demonstration of full-duplex communication on a single laser beam [J]. New Mexico State Univ, 2004, 33(07): 5160–5160.
14. Plett M L, Rabinovich W S, Mahon R, et al. Free-space optical communication link across 16 kilometers over the Chesapeake Bay to a modulated retroreflector array [J]. Optical Engineering, 2008, 47(04):55–76.
15. Öhgren Johan, Kullander Fredrik, Sjöqvist Lars, et al. A high-speed modulated retro-reflector communication link with a transmissive modulator in a Cat's eye optics arrangement [J]. Unmanned/Unattended Sensors and Sensor Networks IV, 2007,24(23): 6736–6736.
16. Rabinovich W S, Moore C I, Mahon R, et al. Free-space optical communications research and demonstrations at the US Naval Research Laboratory [J]. Applied Optics, 2015, 54(31): 189–200.
17. LI Changying, YANG Guowei, BI Meihua, et al. Design and Analysis of Full-Duplex Modulated Retro-Reflector Free Space Laser Communication System [J]. Radio Engineering, 2018 (03): 244–249.
18. Rabinovich W S, Mahon R, Moore C I, et al. Optical scintillation measurements in a desert environment II: retroreflector links [J]. Proceedings of SPIE 8162, 2011, 23 (07): 816–816.
19. Peter G. Goetz, William S. Rabinovich, Rita Mahon, et al. Modulating Retro-Reflector Lasercom Systems for Small Unmanned Vehicles [J]. IEEE Journal on Selected Areas in Communications, 2012, 30(05),443–465.
20. Rosenkrantz Etai, Arnon Shlomi. An innovative modulating retro-reflector for free-space optical communication [J]. Space and Naval Warfare Systems Command, 2013, 2(09):88–136.
21. Thueux Yoann, Avlonitis Nicholas, Erry Gavin. DAZZLE project: UAV to ground communication system using a laser and a modulated retro-reflector [J]. Airbus Group Innovations, 2014, 3(09): 9248–9249.
22. Anonymous. Space Optical Communications Using Laser Beams [J]. NASA Tech Briefs, 2016, 40 (09):45–132.
23. Born Brandon, Hristovski Ilija R, Geoffroy-Gagnon Simon, et al. All-optical retro-modulation for free-space optical communication [J].Optics express, 2018, 26 (04):775–856.
24. Korotkova Olga, Soresi Arianna. Polarization signature of a monostatic double-pass system with a corner-cube reflector in the turbulent air [J]. Applied Optics, 2019, 58(26): 7139–7144.
25. Basel Mounir El Saghir, Mohamed Bakry El Mashade, Ashraf Mohamed Aboshosha. Performance analysis of modulating retro-reflector FSO communication systems over Málaga turbulence channels [J].Optics Communications, 2020, 12 (23):474–474.
26. El Saghir Basel M., El Mashade Mohamed B. Performance of Modulating Retro-Reflector FSO Communication Systems With Nonzero Boresight Pointing Error [J]. IEEE Communications Letters, 2021, 25(06):337–398.
27. Trinh Phuc V, Carrasco Casado Alberto, et al. Experimental Channel Statistics of Drone-to-Ground Retro-Reflected FSO Links With Fine-Tracking Systems [J]. IEEE ACCESS, 2021, 9(12):587–630.
28. Zhang Junling. Preliminary Study on Optical Reverse Modulator Based on Giant Magnetostrictive Material [D]. Chengdu: Chengdu University of Electronic Science and Technology, 2007, 16–19.

29. Zhang Laixian, Sun Huayan, Zhao Yanzhong, et al. Effect of the incidence angle to free space optical communication based on cat-eye modulating retro-reflector [J]. Academy of Equipment (China), 2013, 33(09):8906–8906.
30. Yang Jiguang, Sun Huayan, FAN Guihua. Design of Receiving Amplifier Circuit for Cat'eye Modulated Retro-Reflector Laser Communication [J]. Infrared, 2014, 35(09):31–36.
31. Qiu Hao, Wang Jingyuan, Xu Zhiyong, et al. Research on Optical Communication Technology of Single Light Source Full Duplex Modulated Retro-Reflector [J]. Journal of Optoelectronics·Laser, 2015, 26(12):2314–23110.
32. Xu Shanhe, Xiao Shali, Wang Shan, et al. Research on Modulated Retro-Reflector Optical Communication System Based on Acousto-Optic Modulation [J]. Laser Technique, 2015, 39(05):598–602.
33. Sun Yanrong, Ke Xizheng, Li Yuanhu. Analysis of Factors Affecting the Characteristics of Modulated Retro-Reflector Reflected Light [J]. Infrared and Laser Engineering, 2016, 45(01):222–229.
34. Ke Xizheng, Zhang Han. Indoor Visible Uplink Communication System Based on Modulated Retro-Reflector [J]. Laser & Optoelectronics Progress, 2018, 55(02):115–122.
35. Zhang Han. Design of Visible Light Full-Duplex Communication System [D]. Xian: Xi'an University of Technology, 2018, 25–26.
36. Chen Zhibiao, Lin Hezi. Data Transmission Modeling and Analysis of Modulated Retro-Reflector Laser Communication Link [J]. Laser Journal, 2019, 40(10):160–163.
37. Pan Lu, Wang Jingyuan, Xu Zhiyong, et al. Atmospheric Turbulence and Aperture Average Effect in Modulated Retro-Reflector Optical Communication [J]. Semiconductor Optoelectronics, 2019, 40(01):140–144.
38. Duan Shuaijun, Fan Guihua, Zhang Laixian, et al. Signal-to-Noise Ratio Analysis of Cat'eye Modulated Retro-Reflector Communication Link [J]. Laser and Infrared, 50(11):1384–1389.
39. Li Xiaoyan. Performance of Modulated Retro-Reflector Free-Space Optical Communication under Atmospheric Turbulence [D]. Changchun: Jilin University, 2020, 12–45.
40. Jiayanmei, LU Congmin, Shen Pengfei, et al. Design of Intersatellite Laser Communication System Based on Modulated Retro-Reflector Reflector [J]. Journal of Manned Spaceflight, 2021, 27(06):754–760.
41. Fang Xiaodong, Xu Zhiyong, Wang Jingyuan, et al. Modulated Retro-Reflector Optical Communication Signal Decision Method Based on K-means Algorithm [J]. Journal of Laser Communication Technology, 2021, 45(12):46–410.
42. Chen Tianning, Yang Guowei, Teng Xuyang, et al. Modulated Retro-Reflector Free Space Optical Communication System of Single Light Source Based on CL-IPPM [J]. Optoelectronics · Lasers, 2022, 33(02): 208–216.
43. Xie Mengtong, Wang Jingyuan, Li Jianhua, et al. Analysis of the Influence of Atmospheric Turbulence on Modulating Retro-Reflector Optical Communication [J]. Journal of Modern Optics, 2022, 69(13):236–268.
44. Zhao Honggang. Research on Key Technologies of Modulated Retro-Reflector Laser Communication [D]. Jilin: Changchun University of Science and Technology, 2019, 8–13.
45. Wei Zhenyan, Chen Hui. Design of Receiving Aperture for Optical Wireless Communication Based on Aperture Smoothing Effect [J]. Electronic Design Engineering, 2010, 18(04):55–58.
46. Ren Jianying, Sun Huayan, Zhang Laixian. "Cat's eye" Modulated Retro-Reflector Free Space Optical Communication Technology [J]. Laser and Infrared, 2017, 47(01):98–102.
47. Qiu Hao, Wang Jingyuan, Xu Zhiyong, et al. Modulated Retro-Reflector Optical Communication Technology[J]. Military Communications Technology, 2015, 36(02):29–33.
48. Li Yahong, Fu Yuegang, He Wenjun, et al. Optical design of large field spherical reverse reflector [J]. Acta Photonica Sinica, 2015, 44(11):149–153.
49. Unipixel Displays, Inc. Method and System for Free Space Optical Communication Utilizing a Modulated Electro-Optical Polymer Retro-Reflector. [J]. Telecommunications Weekly, 2014, 12(10):198–216.
50. Duan Shuaijun, Fan Guihua, Zhang Laixian, et al. Analysis of signal-to-noise ratio of communication link based on cat's eye modulating retro-reflector [J]. Laser and Infrared, 50 (11):1384–1388.

51. Li Xiaoyan. Performance of Modulated Retro-Reflector Free-Space Optical Communication under Atmospheric Turbulence [D]. Changchun: Jilin University, 2020.
52. Jiayanmei, LU Congmin, Shen Pengfei, et al. Design of Intersatellite Laser Communication System Based on Modulated Retro-Reflector Reflector [J]. Journal of Manned Spaceflight, 2021, 27(06):754–760.
53. Fang Xiaodong, Xu Zhiyong, Wang Jingyuan, et al. Modulated Retro-Reflector Optical Communication Signal Decision Method Based on K-means Algorithm [J]. Journal of Laser Communication Technology, 2021, 45(12):46–49.
54. Chen Tianning, Yang Guowei, Teng Xuyang, et al. Modulated Retro-Reflector Free Space Optical Communication System of Single Light Source Based on CL-IPPM [J]. Optoelectronics·Lasers, 2022, 33(02):208–216.
55. Xie Mengtong, Wang Jingyuan, Li Jianhua, et al. Analysis of the Influence of Atmospheric Turbulence on Modulating Retro-Reflector Optical Communication [J]. Journal of Modern Optics, 2022, 69(13):236–267.
56. Zhao Honggang. Research on Key Technologies of Modulated Retro-Reflector Laser Communication [D]. Jilin: Changchun University of Science and Technology, 2019.
57. Wei Zhenyan, Chen Hui. Design of Receiving Aperture for Optical Wireless Communication Based on Aperture Smoothing Effect [J]. Electronic Design Engineering, 2010, 18(04):55–57.
58. Ren Jianying, Sun Huayan, Zhang Laixian. "Cat's eye" Modulated Retro-Reflector Free Space Optical Communication Technology [J]. Laser and Infrared, 2017, 47(01):98–102.
59. Zhao Xunjie, Gao Zhiyun, Zhang Yingyuan. Technique of active laser reconnaissance and the applications in the military [J]. Optical Technique, 2003, 4:415–417.
60. Li Zhan. Research on optical reverse modulator with large field of view "Cat's eye" structure [D]. Chengdu: University of Electronic Science and Technology of China, 2008.
61. Wei Guanghui, Yang Peigen. Application of laser technology in ordnance industry [M]. Beijing: Ordnance Industry Press, 1995, 51–66.
62. Zhao Yanzhong, Sun Huayan, Song Fenghua, et al. Effect of focal shift on reflecting character of "cat eye" effect [J]. Laser Technology, 2008, 1:71–74.
63. Qiu Hao, Wang Jingyuan, Xu Zhiyong, et al. Modulated Retro-Reflector Optical Communication Technology [J]. Military Communications Technology, 2015, 36(2):29–33.
64. Zhang Laixian, Sun Huayan, Fan Guihua, et al. Progress in free space optical communication technology based on cat-eye modulating retro-reflector [J]. Chinese Optics, 2013, 6(5): 681–691.
65. Li Yahong, Fu Yuegang, He Wenjun, et al. Optical design of Wide-viewing-angle spherical retroreflector [J]. Acta Photonica Sinica, 2015, 44(11):149–153.
66. G. Yang, C. Li, M. Bi, et al. Performance analysis of full duplex modulating retro-reflector free space optical communications over single and double Gamma-Gamma fading channels [J]. IEEE Transactions on Communications, 2018, 66(8):3597–3609.

Detectors and Their Noise Models

6

Contents

Introduction 158
Photodetectors 158
Photodetector Classification 159
Photodetector Performance Indicators 162
Photodetector Types 164
Source Noises 164
Photomultiplier Tubes 166
Avalanche Photodetector 169
PIN Photodetectors 172
Four-Quadrant Detectors 175
Quantum-Dot Infrared Detectors 179
Balanced Detectors 182
Double Balanced Probes 185
Photodetector Noise 188
Sources of Photodetector Noise 188
Typical Photodetector-Noise Models 190
Summary 197
References 198

Abstract

In optical wireless communication system, photoelectric conversion is realized by photodetector at the receiver, and the noise in the detector will directly affect the photoelectric conversion efficiency. Based on the background of optical wireless communication, according to the generation mechanism of photodetector noise, this chapter describes the shot noise, multiple generation compound noise, 1/f noise, and thermal noise of the detector, in this order, and gives the corresponding mathematical models. The noise models of avalanche photodetector, original negative signature detector, photomultiplier tube, four quadrant detector, quantum dot infrared detector, and balance detector are analyzed, in this order. Finally, the direction of further research in this field is pointed out.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_6

Keywords

Optical wireless communication · Photodetector · Noise model

Introduction

In an optical wireless communication system, the photoelectric signal is converted by a photodetector at the receiving end, and the noise in the detector directly affects the photoelectric-conversion efficiency. In this chapter, the photodetector noise-generation mechanisms are described, and the corresponding mathematical models are presented. The mechanisms include scattered-particle noise, generation-compound noise, 1/f noise, and thermal-detector noise. The noise models of the avalanche photodetector, original negative detector, photomultiplier tube, four-quadrant detector, quantum-dot infrared detector, and balanced detector are analyzed.

Photodetectors are key devices that convert optical signals into electrical signals, and the effectiveness and stability of their output signals are important indicators of their working performance. In this chapter, detectors commonly used in optical wireless communication systems are summarized, including the detector's working principle, performance indicators, application scenarios, and noise models. Finally, the directions and challenges worthy of further study and discussion in this field are discussed.

Photodetectors

An optical wireless communication system includes channel coding and decoding, signal modulation and demodulation, optical amplification and coupling, optical transceiver antennas, and acquisition, alignment, and tracking systems. A laser or light-emitting diode is used as the signal light source, and the signal is modulated after channel coding.

Modulation is divided into active and passive, and active modulation can be divided into direct and indirect. A laser is directly output or amplified by an optical amplifier, which includes semiconductor optical amplifiers and optical-fiber amplifiers. A single or multiple antennas are used for transmission in free space. The transmission channels can be atmospheric, indoor-scattering, ultraviolet-scattering, and underwater.

The receiving end uses a single or multiple antennas for optical-signal reception. The optical-coupling module couples the signal into a single-mode fiber, after preamplification. It uses direct or coherent detection to convert the optical signal to an electrical signal, and then after demodulation and decoding, it delivers the information to the destination. The photoelectric detector is the core of the receiving end.

Photodetector Classification

Photodetectors are a class of devices that can convert optical-radiation signals into electrical signals. Photodetectors commonly used in the field of optical wireless communication can be divided into two categories, based on their physical-mechanism characteristics: photon detectors, based on the photoelectric effect, and thermal detectors, based on the temperature-change effect [1]. Figure 6.1 shows the classifications of some of the photodetectors commonly used in the field of optical wireless communication, based on their external and internal photoelectric effects.

When incident light illuminates the surface of a photoelectric-based detector, a photon of the incident light collides with the electrons in the material to produce new photoelectrons, which overflow from the detector surface, resulting in an external photoelectric effect. If the charge carriers generated inside the detector material do not overflow from the surface, but are bound to the inside of the detector material, it is called an internal photoelectric effect.

The external photoelectric effect generally includes photoconductive, photovoltaic, and magneto-optical effects. The internal photoelectric effect includes the photon-emission effect. The quantum-well and the quantum-dot infrared detectors use the photoconductive effect. For photocells, the positive–intrinsic–negative detector (PIN photodetector) and avalanche photodetector are photovoltaic-effect detectors. Photoresistors belong to detectors using the photoconductive effect. Vacuum photocells, gas-filled photocells, and photomultiplier tubes take advantage of the photon-emission effect.

In a thermal detector, a detector material that absorbs incident light will produce a thermoelectric effect, resulting in a change in resistivity. Commonly used thermal detectors include thermistors and thermopiles.

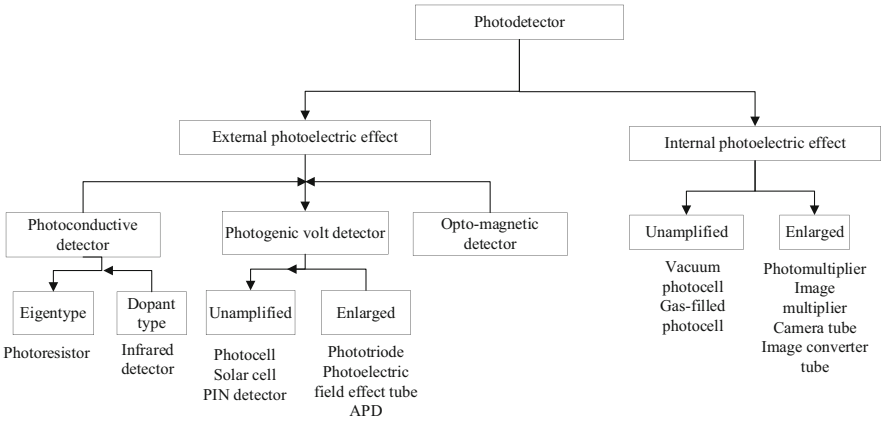


Fig. 6.1 Photodetector classification

Photoconductive Effect

When incident light irradiates the surface of a semiconductor material, the material absorbs energy. At this time, the carrier concentration gradually increases, the conductivity increases, and the resistance decreases; this is the photoconductive effect.

A photoconductive detector, which is a uniform intrinsic semiconductor with a photoconductive effect, can be represented by the model shown in Fig. 6.2. A voltage U is applied to both ends of the semiconductor to form electrodes. The current generated by the circuit can be detected by a galvanometer. When the optical metric is changed, the current of the circuit changes, that is, the resistance of the detector reverses.

Photovoltaic Effect

Figure 6.3 shows a schematic diagram of the photovoltaic effect with the presence of charge carriers in the P-N junction, with two types of carriers, majority and minority. Most carriers will constantly drift in the interior of the P-N junction, forming a self-established electric field E inside the P-N junction. When the P-N junction and the area around it are irradiated with light, a few charge carriers will be present near the P-N junction.

Under the action of the self-built electric field E, the electrons drift to the N region and the holes drift to the P region. The N region, where electrons drift, carries a negative charge. The P region, to which the holes drift, has a positive charge. The electric field generated by the control of a few carriers is an additional electric field, that is, the photogenerated electric field.

When a reverse voltage is applied to the P-N junction, the resulting reverse current gradually increases as the reverse voltage increases. The resulting photoelectric signal can be observed using photocurrent I.

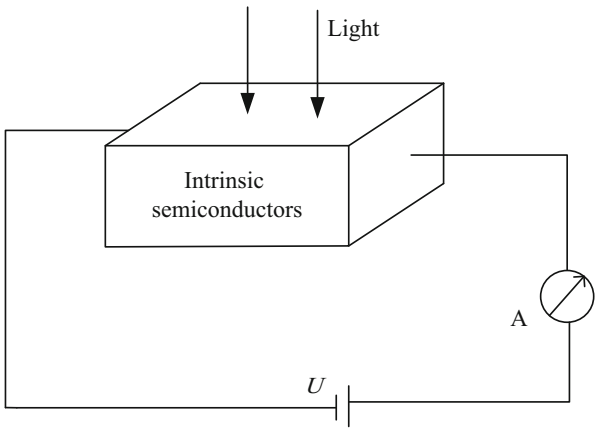


Fig. 6.2 Simple model of a photoconductive detector

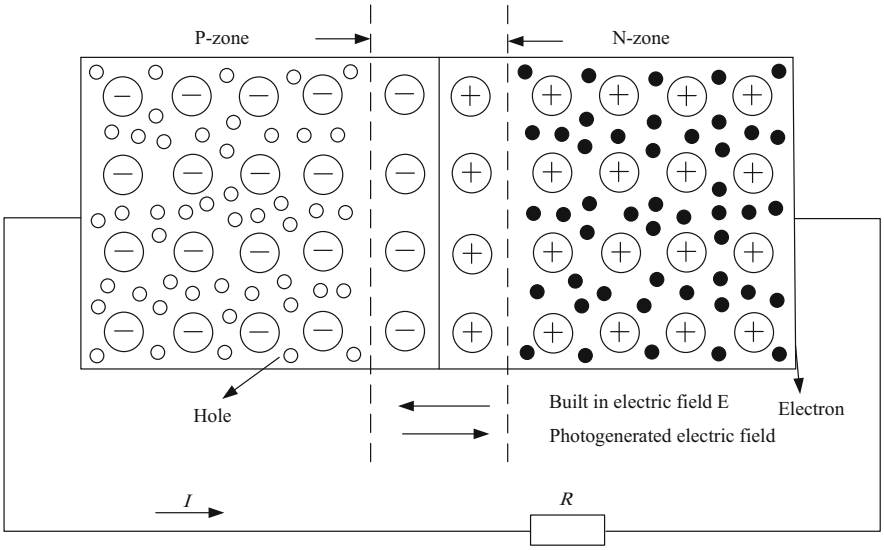


Fig. 6.3 Photogenerated voltage effect of a P-N junction

Photoelectron-Emission Effect

According to the quantum theory of light, when incident light strikes the surface of the device material, if the frequency of the incident light increases, the energy absorbed by the electron increases, and part of the energy is used for collision loss. If the remaining energy is sufficient to overcome the barrier on the surface of the material, the electrons overcoming the barrier will overflow from the surface of the material and enter the space environment.

The photoelectron energy-level transitions are shown in Fig. 6.4. The maximum kinetic energy of the photoelectrons that overflow the surface of the material is described by Einstein's equation:

$$E_k = hv - w, \tag{6.1}$$

where w is the work function of the material in joules, h is Planck's constant, and $E_k = 1/2^{mv^2}$ is the kinetic energy of a photoelectron spilling over the surface of the material. The mass of the photoelectron is represented by m, and its velocity is represented by v.

When the incident light irradiates the surface of the material, the kinetic energy of the photoelectrons does not change with the change in light intensity but is only related to the frequency of the incident light. When the incident-light frequency increases, the kinetic energy increases; when the incident-light frequency decreases, the kinetic energy decreases. When photoelectrons leak from the surface of the material, there is a critical threshold value, that is, when the kinetic energy is 0, the photoelectrons simply reach the surface of the material. In this case, the frequency of the incident light is the limiting frequency.

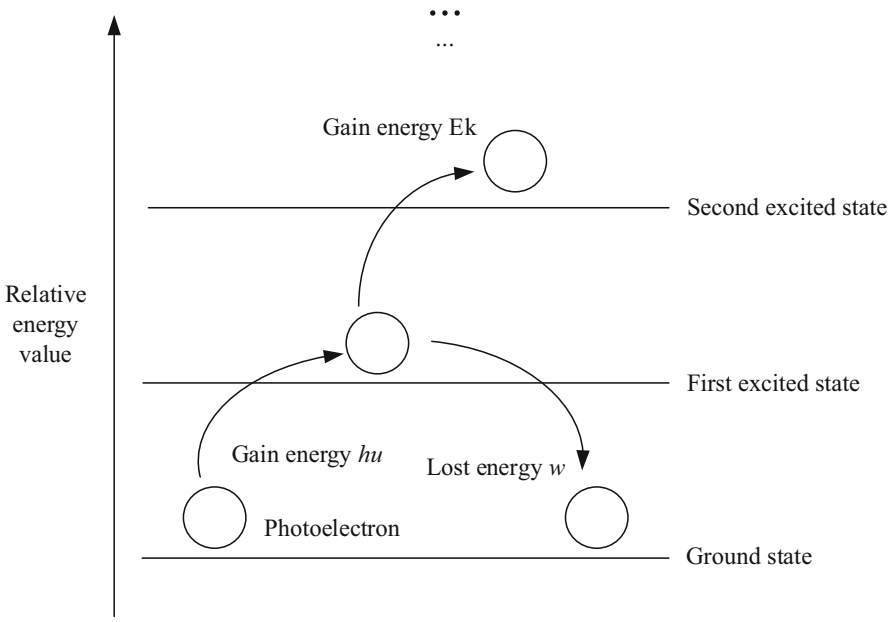


Fig. 6.4 Photoelectron energy-level transition diagram

Photodetector Performance Indicators

Quantum Efficiency

The quantum efficiency is defined as the average number of electrons released by an incident photon, η, and is expressed as

$$\eta = \frac{Ih\nu}{Pe}, \tag{6.2}$$

where η is the quantum efficiency, I is the average value of the photocurrent, e is the electron charge, P is the optical power, h is Planck's constant, and ν is the frequency of the incident light. Detectors with higher quantum efficiencies have better performance.

Noise-Equivalent Power

The noise-equivalent power (NEP) reflects the detection capability of the detector and is generally described under the premise of a unit signal-to-noise ratio, that is,

$$\text{NEP} = \frac{PV_s}{V_n}, \tag{6.3}$$

where NEP is the noise-equivalent power, V_s is the voltage intensity of the signal, V_n is the voltage intensity of the noise, and P is the incident light power. The larger the

NEP, the weaker the detection ability; the smaller the NEP, the stronger the detection ability.

Responsivity

The ratio of the output optical power to the input optical power of the detector is the responsivity R, which can be subdivided into voltage and current responsivities. The voltage responsivity R_u is defined as the ratio of the output voltage to the incident light power, that is,

$$R_u = \frac{U_s}{P}, \tag{6.4}$$

where R_u is the voltage responsivity, U_s is the output voltage of the photodetector, and P is the incident-light power.

The current responsivity R_i is defined as the ratio of the output current of the photodetector to the incident-light power, as follows:

$$R_i = \frac{I_s}{P}, \tag{6.5}$$

where R_i is the current responsivity, I_s is the output current of the photodetector, and P is the incident-light power.

Spectral Responsivity

The spectral responsivity $R(\lambda)$ is defined as the ratio of the output voltage or current of the detector to the incident light power, when it is illuminated by monochromatic light of wavelength λ.

$$R_V(\lambda) = \frac{V_S(\lambda)}{P(\lambda)} \tag{6.6}$$

$$R_I(\lambda) = \frac{I_S(\lambda)}{P(\lambda)}, \tag{6.7}$$

where $R_V(\lambda)$ and $R_I(\lambda)$ are the spectral responsivities, $V_S(\lambda)$ is the output voltage, $I_S(\lambda)$ is the output current, and $P(\lambda)$ is the power of the incident light.

Frequency Responsivity

The frequency responsivity $R(f)$ is affected by the incident-light frequency and is defined as

$$R(f_\tau) = \frac{R_0}{\left[1 + (2\pi f_\tau)^2\right]^{1/2}}, \tag{6.8}$$

where $R(f_\tau)$ is the responsivity when the frequency is f_τ, and R_0 is the responsivity when the frequency is 0. The time constant τ of the detector is a fixed value, usually determined by the material of the detector and the external circuit.

Detectivity and Normalized Geochemical-Exploration Measures

The detectivity D reflects the sensitivity of the detector and is the reciprocal of NEP:

$$D = \frac{1}{\text{NEP}}, \tag{6.9}$$

where D is the detection degree and NEP is the noise-equivalent power of the detector. The normalized geochemical exploration measure D^* is defined as follows:

$$D^* = D \cdot (A_d \Delta f)^{1/2}, \tag{6.10}$$

where D^* is the normalized geochemical-exploration measure, D is the detectivity, A_d is the area of the detector, and Δf is the amplifier bandwidth.

Photodetector Types

Source Noises

According to the different emission and modulation mechanisms of the system, source noise can be divided into laser light source noise and noise in LED-modulated signal light in visible light communication using LED as the light source. When the communication signal modulates the laser emission, the source noise can be roughly divided into two parts: the inherent thermal noise of the laser itself, and the distortion of the laser emission compared to the modulated signal. The study of the above noise/distortion of lasers has started since the invention of lasers in the 1960s. In 1966, Pauwels derived the series expressions for the amplitude and phase noise of lasers from the Van del Pol equation of quantum mechanics, but their expressions were too complex [2]. In the same year, Leeson obtained the spectral approximation to 1/f noise from a simpler spectral model [3]. In 1970, Ziel investigated the scattering grain noise model for GaAs lasers [4], and the results were widely quoted. In 1982, Harder et al. [5] solved the semiconductor rate equation to obtain the light intensity and spectral undulation of a semiconductor laser and obtained the laser's general equivalent circuit of the lase. In 1989, Petermann gave an exhaustive summary of the various types of noise in lasers [6]. Due to the complex form of the solution of the rate equation, it is often necessary to describe it by the series method, and this signal nonlinear distortion is also a kind of noise affecting the communication system. In 1992, Kuo conducted theoretical analysis and experimental research on the influence of gain saturation, relaxation oscillation, spatial hole burning, and other factors on the nonlinear noise of LD [7]. In 1994, Bihan et al. [8] used perturbation method and Bessel function method to solve the rate equation

and derived the expression of LD nonlinear intermodulation distortion, and the results show that the FM signal distortion is more obvious than the AM signal. In 1997, Chai et al. [9] analyzed the effect of phase noise of semiconductor laser on the characteristics of high-speed optical fiber communication system, the results show that the dispersion characteristics of optical fiber make the phase noise of laser will be transformed into intensity noise at the receiving end, and the sensitivity of the optical receiver deteriorates and shows the "saturation" phenomenon on the BER curve. In 2002, Fukuda et al. [10] elucidated experimentally and theoretically the reason for the increase of the residual spectral linewidth during the degradation of the device, based on the multi-quantum-well-distributed feedback lasers, and proposed a correlation between the 1/f noise and the degradation of the semiconductor lasers. In the same year, Lawrence conducted a detailed study on nonlinear distortion in LDs driven by FM signals and showed that factors such as the modulation regime, the modulation frequency, the magnitude of the injected current, and the length of the external cavity affect the nonlinear dynamic distortion of the lasers [11].

In 2006, Yu et al. [12] designed a new type of low-noise fiber laser using a composite cavity structure, which effectively suppressed the intensity noise of the light source. In 2010, Domenico et al. [13] proposed an approximate formula to handle the relationship between laser frequency noise spectrum and laser linear, which can be applied to any noise spectral density. In 2012, Liu et al. [14] studied the effects of laser phase noise on modulation format selection, parameter optimization, and carrier phase estimation performance in coherent fiber optic communication systems. Simulation results showed that when the carrier phase estimation method is used in coherent fiber optic communication systems, the optimal average length of the symbol sequence in the algorithm is inversely proportional to the laser phase noise. In 2015, Mahmoud conducted a detailed study on the impact of LD noise and distortion in Radio Over Fiber on the system, pointing out that intermodulation distortion is more severe than harmonic distortion during laser modulation [15]. In 2016, Chen et al. [16] derived the expression of nonlinear intermodulation distortion in lasers using the Bessel function, studied the impact of nonlinear intermodulation distortion on the carrier intermodulation noise ratio and bit error rate performance of the system, and provided the selection of the optimal modulation index of the system and the corresponding relationship with the nonlinear coefficient. In 2016, Chen proposed a novel nonoptical heterodyne detection method based on subcarrier modulation and conducted system simulation experiments under additive noise, multiplicative noise, and mixed noise, providing experimental basis for the application of this nonoptical heterodyne detection method [17]. In 2017, Song et al. [18] qualitatively analyzed the phase noise of fiber lasers with different pump powers, seed source powers, and wavelengths using power spectra and integral spectra. The results showed that the higher the pump power, the greater the proportion of low-frequency phase noise; for different seed source powers, the proportion of phase noise in the test frequency band is basically at the same level; and the proportion of phase noise in the high-frequency range varies for different seed wavelengths tested, with the highest proportion occurring when the seed wavelength is 1560.48 nm. In 2019, Bai derived the relationship between laser linewidth,

frequency noise power spectral density, and phase noise power spectral density. A method for calculating laser linewidth by measuring laser phase noise was obtained, and numerical simulation and characteristic analysis were conducted on the output laser spectral line of the laser [19]. In 2022, Du et al. [20] proposed and conducted a theoretical study aimed at exploring the noise codriving chaotic synchronization phenomenon based on a three-stage-distributed Bragg reflection semiconductor laser.

Photomultiplier Tubes

Working Principle

The principle of the photomultiplier tube is illustrated in Fig. 6.5. The electron-focusing system includes a focusing pole D, photocathode K, and anode a, from which the incident light enters into the interior of the photomultiplier tube. Under the action of the photocathode, the incident light will converge to form a beam of light that strikes the first multiplier pole. Under the excitation of the first multiplier pole, the reflected beam produces a large number of secondary electrons, which are driven by an electric field to the second multiplier pole, producing more secondary electrons; this continues to D_{10}.The last incident beam passes through D_{10}. The pole multiplication reaches the anode, and the output photocurrent at load $R_L V$ is formed on V_0.

Research Progress

In 1972, Shun et al. [21] developed several new types of photocathode materials and analyzed the corresponding processing technology. A new secondary-emission technology was used in the photomultiplier, which improved its gain characteristics.

In 1989, Li [22] discovered how the polarization effects of the grating and photomultiplier tubes varied with the wavelength. The experimental results showed

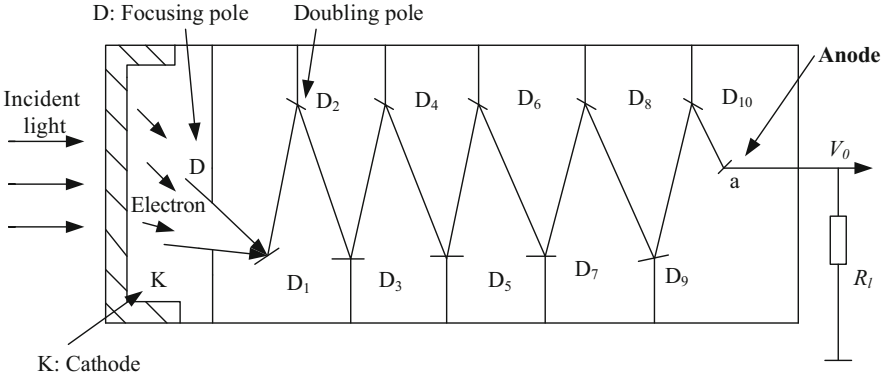


Fig. 6.5 Schematic diagram of a photomultiplier tube

that the polarization effect of the grating was greatly affected by the change in wavelength, in the range of 420–740 nm. Therefore, the polarization effect of the grating must be considered when measuring the polarization.

In 2000, Chow [23] analyzed basic photomultiplier-control technology and its improvements. Through the continuous improvement of basic devices, a photomultiplier tube can be used as a weak light-detection device.

In 2004, Guo et al. [24] discussed the noise sources of photomultiplier tubes, including loose-particle, thermal, generation-compound, and 1/f noise, and established the corresponding noise model.

In 2014, Chowdhuri et al. [25] analyzed the spectral detection based on a photomultiplier array; the detection method had a fast time response of 10 μs and a spatial resolution of 3 cm.

In 2015, Maltseva et al. [26] used photomultiplier tubes to detect Cherenkov light; the detector's spatial resolution reached 3 m for 20 m long fibers. In addition, the resolution could be improved by optimizing the parameters of the optical fiber and photomultiplier, and the spatial resolution of the detector reached 0.5 m.

In 2016, Chen et al. [27] analyzed the performance of a silicon photomultiplier tube, designed a nuclear-radiation detector, developed a radiation-measurement probe, and tested its radiation performance. This detector has some reference value for the design of nuclear-radiation detectors.

In 2017, Zhao et al. [28] developed a one-dimensional single-photon position-sensitive silicon photomultiplier tube, with a single-photon position resolution of 393.4 μm. It improved from 393.4 μm to 56.2 μm when the number of photoelectrons was increased from one to seven.

In 2018, Takahashi et al. [29] analyzed the detection principle of low-intensity light and estimated the average signal generated by a single photoelectron, using the fact that photoelectrons conform to the Poisson distribution.

In 2019, Elorrieta et al. [30] described the detailed characteristics of a 3-inch photomultiplier, including the dark count rate, probability of associated pulses, timing characteristics, gain, and temperature dependence of the peaks and valleys. The effect of conductive objects near the photocathode on photomultiplier noise was also shown.

In 2020, Sun et al. [31] designed a system consisting of ultraviolet (UV) light and highly sensitive photomultiplier tubes to simulate very cloudy harbor water. It transmitted data at a rate of 85 Mbit/s using on–off control modulation and established a non-direct-looking link with a transmission distance of 30 cm.

In 2021, Ning et al. [32] investigated the performance of devices under various light intensities and photomultiplier-tube gains and proposed a small-volume receiver-control system and related adaptive-control strategy to improve stability and reliability.

Table 6.1 summarizes the relevant properties of photomultiplier tubes. Photomultiplier tubes have applications in the ultraviolet communication field, as well as weak optical-signal detection.

Table 6.1 Research progress of photomultiplier tubes

Researcher	Year	Development
Shun	1972	Improved the gain characteristics of photomultiplier tubes
Li	1989	Grating polarization effect must be considered in polarization measurement
Zhou	2000	Photomultiplier tubes can be used as weak light-detection devices
Guo	2004	Discussed the sources of noise and built a noise model
Maltseva	2015	Detector spatial resolution reached 3 m
Chen	2016	Developed nuclear-radiation detector
Zhao	2017	Single-photon position resolution is 393.4 μm
Takahashi	2018	Conversion of photons to photoelectrons
Elorrieta	2019	Effect of conductive objects near the photocathode
Sun	2020	Non-direct-looking link with a data rate of 85 Mbit/s
Ning	2021	Small-volume receiver-control system

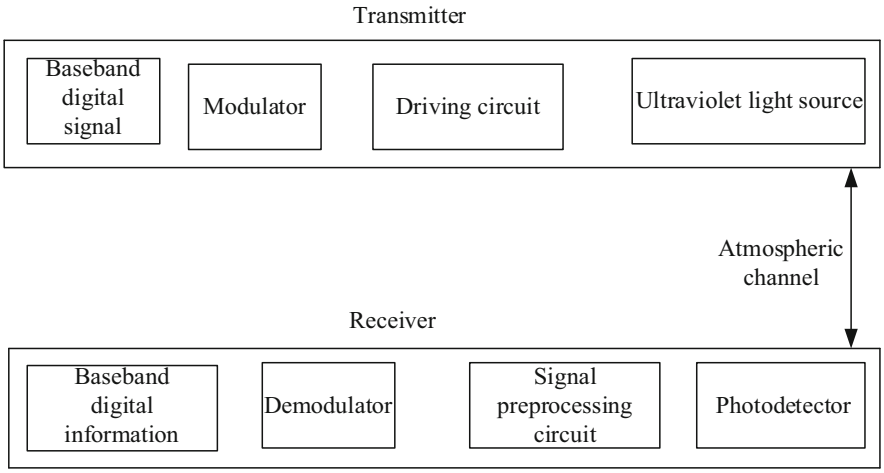


Fig. 6.6 Structure of a UV communication system [33]

Typical Application Scenarios

Figure 6.6 shows the structure of a UV communication system that consists of transmitting, receiving, and atmospheric channels [33]. The transmitting component modulates the original electrical signal and transforms it into a signal suitable for transmission in the channel. The signal is disrupted by interference in the atmospheric channel, and the receiving part must recover the original signal from the received disrupted signal.

The receiving part is composed of an ultraviolet photomultiplier tube, preprocessing circuit, demodulation module, and beacon module. The ultraviolet detector collects the ultraviolet signal and converts the corresponding optical signal into an electrical signal. The preprocessing circuit amplifies and filters the electrical signal, and the demodulation module demodulates it and converts it into a baseband

digital signal. The UV photomultiplier is currently the most ideal UV detector. It has been widely used for nondirect UV communication because of its large detection area, high gain, and low dark current.

Avalanche Photodetector

Working Principle

An avalanche photodetector is based on an avalanche photodiode (APD). Figure 6.7 shows a schematic diagram of the avalanche breakdown. When incident light illuminates the diode surface, the light is absorbed by the P-N junction to form a photocurrent. The photocurrent is affected by the reverse bias and increases with the increase in reverse bias. When the reverse bias is continuously increased, the photocurrent exponentially increases. At this time, the so-called "avalanche" phenomenon occurs.

Research Progress

In 1980, Yang et al. [34] studied avalanche-photodiode technology and proposed a new avalanche photodiode that had high gain and low noise and could be used for long-wavelength detection.

In 1990, Watkins et al. [35] observed the characteristics of laser-induced electrical failures in biased silicon avalanche photodiodes. When monitoring current–voltage characteristic changes, two types of changes were observed: The first was a large increase in leakage current, which can be modeled by introducing defects into the depletion region through deep melting transients. The second type was a catastrophic failure, in which an electrical short circuit occurred in the device after irradiation; it can be modeled by excessive current density in the photodiode junction.

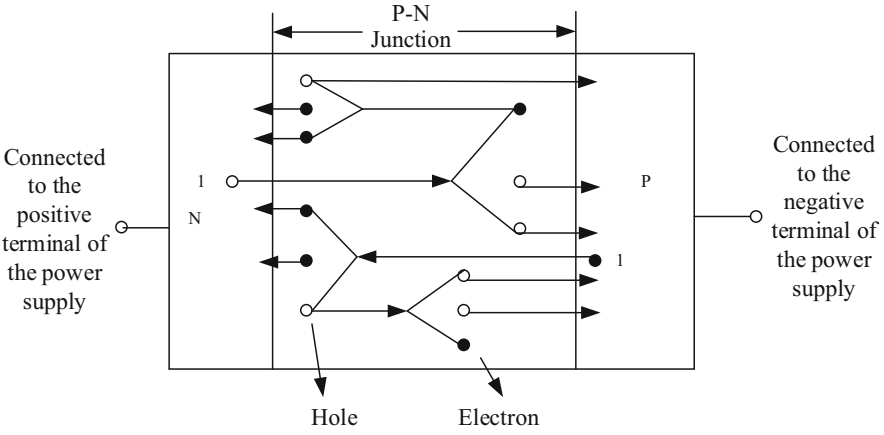


Fig. 6.7 Schematic diagram of the avalanche breakdown

In 1996, Li et al. [36] designed a superlattice avalanche photodetector. The optimal structural parameters were obtained as follows: The thickness of the i-Si avalanche zone was 1.8 μm–2 μm. The carrier concentration in the p-Si zone was 1.018 cm^{-3}, and the thickness was 17 nm. The total thickness of the superlattice was 340 nm. The detector could operate in the range of 1.3 μm–1.6 μm.

In 2005, Ng et al. [37] concluded that collision ionization in the InGaAs absorption layer increased the excess noise in the InP avalanche photodiode. Moreover, strict limits needed to be imposed on the doping in the charge-control layer if the avalanche noise was to be minimized.

In 2009, Guo et al. [38] used passive-suppression technology to determine that dark current does not reflect Pass-through characteristics, which is an obvious difference between photocurrent and dark current. Through the above research, it was concluded that to improve the signal-to-noise ratio of a single-photon detector, the through-through characteristics could be used to properly select the reverse-bias voltage of an avalanche photodiode in Geiger mode.

In 2010, Chitnis et al. [39] investigated a compact readout circuit that could be integrated within a single single-photon avalanche detector. The avalanche process in photodetectors can be suppressed by using the smallest transistor size.

In 2011, Zhang et al. [40] analyzed the scattered-particle noise and 1/f noise of avalanche photodiodes and established their signal-to-noise ratio (SNR) model.

In 2012, Guo et al. [41] studied the relationship between dark noise and the detector temperature. The results showed that when the temperature was reduced from 60 °C to −40 °C, the noise-equivalent power decreased by one order of magnitude. If the cooling, spatial filtering, and spectral filtering methods were combined, the sensitivity of the detector increased by 7.7 times.

In 2013, Adamo et al. [42] analyzed novel silicon photomultiplier tubes manufactured on silicon p-type substrates; they could be used as sensitive power meters for low-photon fluxes.

In 2014, Youn et al. [43] studied an 850 nm photoelectric integrated-circuit receiver and found that the signal-to-noise characteristics of avalanche photodiodes depended on the reverse-bias voltage.

In 2015, Chen et al. [44] studied a low-voltage germanium-waveguide avalanche photodetector with a gain × bandwidth product higher than 100 GHz and developed a photon receiver based on this photodetector. Using the bit-error rate measurement, the avalanche gain was 6 and the sensitivity was improved by 5.8 dB at a −5.9 V bias.

In 2016, Goykhman et al. [45] studied a silicon plasma Schottky photodetector and achieved a responsivity of 0.37 A/W with a reverse bias of 3 V.

In 2017, Yin et al. [46] simulated an avalanche photodetector using benzocyclobutene as a bonding-layer material, with a detection efficiency of 96.7%.

In 2018, Jukic et al. [47] discovered that the excess noise of avalanche photodiodes was related to the avalanche gain. By analyzing the structure of the avalanche photodiode, it was found that this correlation could be suppressed by modulation doping.

In 2019, Huffaker [48] proposed a new avalanche photodiode consisting of a vertical InGaAs array that could be used for single-photon detection, greatly reducing the avalanche volume and number of filling traps.

In 2020, Zhou et al. [49] studied a linear array of 1 × 64 highly uniform 4H-SiC avalanche photodiodes for UV detection; it achieved a highly uniform breakdown voltage and a fluctuation of less than 0.5 V.

In 2021, Eid et al. [50] studied the pulse broadening and signal-fiber bandwidth of a silicon-doped germanium-fiber link in terms of the temperature variation and found that the bit-error rate of the silicon avalanche photodetector at the receiving point was related to the temperature.

In 2022, Li et al. [51] studied the excess noise of avalanche photodiodes, analyzed two noise-measuring methods, including the direct-power measurement method and phase-sensitive detection method, and also analyzed and discussed their advantages and disadvantages. On this basis, three methods of reducing the excess noise were obtained: selecting materials with a low impact-ionization coefficient ratio, reducing the thickness of the multiplier layer, and using APD impact-ionization engineering.

Table 6.2 shows the development of avalanche photodiodes. The gain and sensitivity of the detector are discussed in the above literature. A noise model of the avalanche photodetector was established by analyzing the source of the noise, noise-suppression techniques were studied, and the InGaAs/InAlAs detector as a mature material was further analyzed.

Table 6.2 Research progress on avalanche photodiodes

Researcher	Year	Development
Yang	1980	APDs can be used for long-wavelength detection
Watkins	1990	Observed the characteristics of laser-induced electrical failures
Li	1996	Superlattice avalanche photodetector is studied and optimized
Ng	2005	Strict restrictions on the doping of charge-control layers
Guo	2009	Proper selection of the reverse bias of an avalanche photodiode
Chitnis	2010	Avalanche process in the photodetector is suppressed
Zhang	2011	SNR models of granular noise and 1/f noise are established
Guo	2012	Sensitivity of the detector is improved by 7.7 times
Adamo	2013	Used as a sensitive power meter for low photon flux
Youn	2014	SNR depends on the reverse bias voltage
Chen	2015	Sensitivity is improved by 5.8 dB
Goykhman	2016	Responsivity of 0.37 A/W was achieved
Yin D	2017	Detection efficiency reached 96.7%
Jukic	2018	Excess noise can be suppressed by modulation doping
Huffaker	2019	Vertical InGaAs arrays can be used for single-photon detection
Zhou	2010	High uniform breakdown voltage is achieved
Eid	2021	Bit-error rate is temperature-dependent
Li	2022	Excess noise of avalanche photodiodes is studied

Typical Application Scenarios

Figure 6.8 shows the principle block diagram of a detection and receiving system. A weak optical signal that is diffusely reflected back is converted into an electrical signal by an avalanche photodetector and amplified by an amplifier. Then, it enters a comparator for comparison. The bias control circuit provides the avalanche photodetector with the corresponding bias voltage, giving the detector an optimal multiplier factor under the optimal bias voltage. The voltage amplitude control adjusts the output-signal amplitude. Adjusting to the optimal amplitude can reduce the blind area of the ranging system. The threshold-circuit comparison voltage is discriminated to output a useful signal level. The comparison voltage must be simultaneously lower than the signal voltage and higher than the noise voltage. Finally, further processing is conducted using the oscilloscope.

PIN Photodetectors

Working Principle

The PIN detector separates the p-type region from the n-type region by adding a lightly doped thin layer in the middle of the P-N junction. The structural diagram of the PIN diode is shown in Fig. 6.9. The junction capacitance of the PIN detector is very small, and the depletion-layer width of the P-N junction is directly related to the

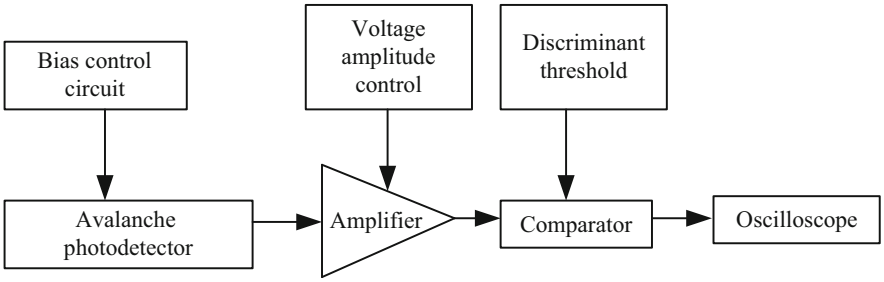


Fig. 6.8 Block diagram of the receiving system

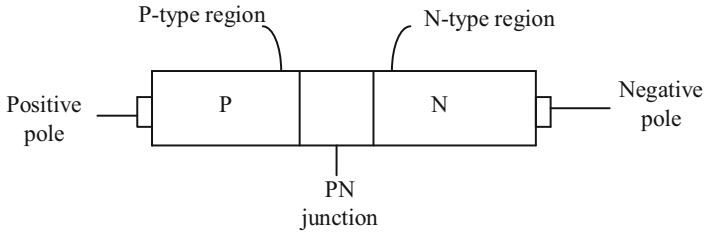


Fig. 6.9 Structural diagram of a PIN diode

reverse bias. When the reverse bias increases or decreases, the depletion-layer width increases or decreases correspondingly. The PIN detector structure is simple.

Research Progress

In 1986, Shu B researchers [52] successfully manufactured a silicon PIN photodetector based on high-resistance monocrystalline silicon using a mature semiconductor planar process.

In 1992, Zhang et al. [53] developed a state-space model for PIN photodetectors. This model produced the detector's time-domain and frequency-domain responses to any form of optical excitation, including terminal, uniform, and exponential illumination. Transmission time limits and circuit RC time constants were included in the model.

In 1997, Zhang et al. [54] used a coaxial package for the InGaAs PIN photodetector chip to measure the C–V characteristics and transient response of the coaxial package device. The results showed that the transient response of the device was greatly improved; the rise time was reduced from 85 ps to less than 25 ps, the half-height width was reduced from 210 ps to 85 ps, the capacitance was reduced by approximately 0.4 pF, and the bandwidth was increased to more than 6 GHz.

In 2004, Jutzi et al. [55] investigated a SiGe PIN photodetector and achieved an optical responsiveness of 160 mA/W at 1300 nm.

In 2010, Giannopolous et al. [56] designed a vertical-cavity surface-emitting laser that was integrated with a PIN photodetector monolith for optical-position sensing. The design was integrated into a silicon platform to form a microsystem. By using a metal grating as a position meter, the sensor microsystem could measure the difference in the reflected power of the grating as it traveled parallel to the sensor.

In 2011, Shao et al. [57] studied the growth and preparation of a back-lit AlGaN/GaN PIN solar blind UV photoelectron detector. The peak responsivity of the photodetector was about 0.055 A/W at 270 nm at a 2.5 V reverse bias, and the dark-current measurement of a 1500 μm diameter device was approximately 25 pA at a 0.3 V bias.

In 2012, Guan et al. [58] analyzed various photoelectric-conversion circuit noises where the PIN detector was located and proposed a method to reduce the noise in the photoelectric conversion circuit to improve the signal-to-noise ratio and measurement accuracy.

In 2014, Wu et al. [59] studied the threshold characteristics of a PIN detector in optical wireless communication. The experiment showed that the detector reached the depth saturation at a 1064 nm wavelength, resulting in serious intercode crosstalk; thus, it could not meet the communication indicators.

In 2015, Chaudhari et al. [60] proposed a low-cost, low-temperature process with precise-thickness films, which was suitable for the low-cost industrial manufacturing of PIN-based detectors.

In 2016, Rouse et al. [61] manufactured a Ge-based PIN detection device on 300 mm silicon wafers using high-throughput, large-area complementary metal-oxide semiconductor technology.

In 2018, Xu et al. [62] used a laser light source with a 10 ns pulse width and a PIN detector for photoelectric conversion. When the applied bias voltage was 89.449 V, the output current signal needed to be amplified by a transresistance amplifier, and the amplified voltage amplitude reached −3.7 V.

In 2019, Shu et al. [63] studied the influence of the main noise sources of a 10 Gbit/s PIN photodetector circuit on the output current signal using an adaptive-filtering algorithm and a peak-pulse median removal algorithm. The results showed that the above denoising method could effectively reduce the measurement error of the output current, improve the measurement accuracy, and increase the stability.

In 2021, Wang et al. [64] designed a highly responsive silicon-based PIN detector with a 4.1 mm × 13.8 mm surface area and a 420 μm thickness. The photocurrent signal was extracted through a ring aluminum electrode, and the detector had a response of 0.6 A/W at a constant laser light source with an 860 nm wavelength.

In 2022, Xiao et al. [65] designed a high-speed monolithic PIN photodetector based on InP and integrated it with a multi-quantum-well semiconductor optical amplifier. The chip allowed for the separate design of the multi-quantum-well semiconductor optical amplifier and PIN photodetector. The PIN photodetector was manufactured with a 300 pA dark current at 3 V.

Table 6.3 illustrates the development of the PIN detector. The above literature mainly discusses the industrial structure and preparation of PIN detectors and their corresponding noise models, as well as the related applications of PIN detectors in optical wireless communication.

Typical Application Scenarios

Owing to the small dynamic range of traditional optical-power detection and the errors in amplification and switching, a weak optical signal is amplified according to

Table 6.3 Research progress of PIN photodetectors (PD)

Researcher	Year	Development
Zhang	1992	Established a state-space model of the PIN photodetector
Zhang	1997	Measured the transient response of a PIN PD with a coaxial package
Jutzi	2004	Optical response of 160 mA/W is achieved at 1300 nm
Giannopolous	2010	PIN photodetector monolithic set created
Shao	2011	Peak responsivity is about 0.055 A/W at 270 nm
Guan	2012	Noise of photoelectric conversion circuit is reduced
Wu	2014	Observed the threshold characteristics of a PIN detector
Chaudhari	2015	Low-cost industrial manufacturing of PIN detectors
Rouse	2016	Ge-based PIN detection devices were manufactured
Xu	2018	Voltage amplitude reached −3.7 V
Shu	2019	Reduced the measurement error of the output current
Wang	2021	Responsivity of the detector at a constant laser source
Xiao	2022	Dark current is 300 pA at 3 V

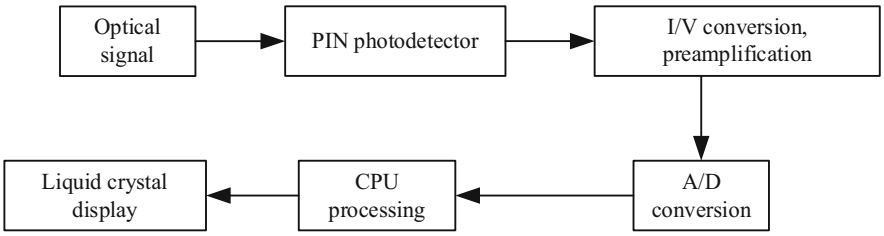


Fig. 6.10 Schematic diagram of an optical power meter

the response characteristics of a PIN photodiode, and a single-chip microcomputer is used for data-correction processing. When it is finally combined with liquid-crystal display (LCD) technology, an optical power-detection meter can be produced.

The basic working-principle block diagram of the optical power meter is shown in Fig. 6.10. The optical signal first passes through the PIN photodetector, which converts the optical signal into an electrical signal; then, the electrical signal goes through an I/V transformation and preamplification processing. The analog signal obtained after processing is converted to digital; the digital signal represents the power size. Then, the CPU processes and analyzes the data. Finally, the power size is displayed on the LCD.

Four-Quadrant Detectors

Working Principle

Following Cartesian-coordinate requirements, diodes with the same performance are placed in the four quadrants, and the detector formed by such a combination is called a four-quadrant detector. A very narrow interval exists between the four quadrants, which is called the dead zone. The detector's detection ability is affected by the dead zone; therefore, the dead-zone area should be reduced as much as possible when making the detector.

Figure 6.11 shows the surface structure of the four-quadrant photodetector. The incident light irradiates the surface of the circular area in the figure, which is called the sensitive surface. The sensitive surface is divided into four parts, namely, A, B, C, and D, and the corresponding areas of these four parts are S_A, S_B, S_C, and S_D, respectively. The corresponding electrodes inside each of these four parts will generate impedance currents I_A, I_B, I_C, and I_D, respectively.

As the incident beam moves, the spot on the detector's surface moves with it, so the true area of light in each quadrant changes, as does the intensity of the light received. The lateral and longitudinal offsets of the light can be determined by the amount of current generated. The formulas for determining the spot position by the area of light received are as follows:

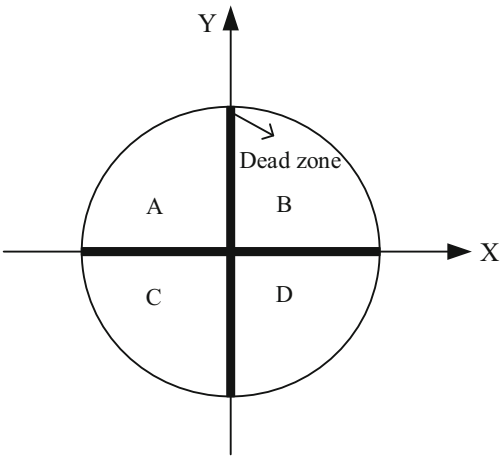


Fig. 6.11 Four-quadrant detector surface-structure diagram

$$\sigma_{\mathrm{x}} = \frac{(S_A + S_D) - (S_B + S_C)}{S_A + S_B + S_C + S_D} = \frac{(I_A + I_D) - (I_B + I_C)}{I_A + I_B + I_C + I_D} \tag{6.11}$$

$$\sigma_y = \frac{(S_A + S_B) - (S_C + S_D)}{S_A + S_B + S_C + S_D} = \frac{(I_A + I_B) - (I_V + I_D)}{I_A + I_B + I_C + I_D}, \tag{6.12}$$

where S_A, S_B, S_C, and S_D are, respectively, the corresponding spot areas of the four quadrants; I_A, I_B, I_C, and I_D are, respectively, the impedance currents generated by the corresponding four quadrant electrodes. σ_x and σ_y are the calculated values for the X and Y axes, respectively, which can be used to analyze the position information of the light spot on the target plane.

Research Progress

In 1979, researchers [66] designed a quadrant detector with a new structure. A P-N junction was used for isolation on the same silicon chip. In addition, a quadrant detector was successfully developed using a four-quadrant patterned oxidation-mask diffusion process.

In 1996, Feng et al. [67] discussed some of the key factors involved in optical roaming using a four-quadrant photodetector, such as the error-data extraction and target-spot size on the active quadrant of the photodetector, in search of more flexible software and hardware design methods.

In 2002, Wang et al. [68] proposed a method to improve the detection accuracy of the high-point position. A Gaussian spot was used as the incident-light model to analyze the solution relationship between the actual position of the spot and the output signal of the four-quadrant detector. Then, the output current signal of the detector was amplified through an analog circuit, and the output voltage signal with noise was digitally filtered. The results showed that the maximum spot-position detection error of this method was only 0.0277 mm, and the root mean square error

was 0.0065 mm, when the four-quadrant detector was in a high-background noise environment.

In 2004, Kuang et al. [69] determined the offset of the two directions of the light spot by measuring the change in the position of its centroid from the laser beam.

In 2005, Chi et al. [70] analyzed the noise of a four-quadrant infrared detector, including loose-particle noise, thermal noise, generation-compound noise, 1/f noise, and the cause of the noise formation.

In 2010, Xin et al. [71] analyzed the factors affecting the tracking accuracy of a four-quadrant detector, including spot size, spot location, and detector noise. The analysis results showed that when using the commonly used four-quadrant centroid algorithm, it was necessary to adopt an appropriate optical system and signal-detection circuit to achieve high-precision tracking-error detection.

In 2013, Yang et al. [72] made a detailed analysis of the dead zone of a four-quadrant detector and showed the influence of the dead zone on the detection sensitivity; the larger the dead zone area, the greater the impact.

In 2014, Schalkwyk et al. [73] made a forward-looking intrinsic day-blind UV four-quadrant Schottky detector based on AlGaN with a Schottky barrier height of 1.22 ± 0.07 eV.

In 2016, Tian et al. [74] proposed an optical-system design method that controlled the spot-size design on a four-quadrant detector by adjusting the entry/exit pupil ratio of the nonfocal system, according to the characteristics of inter-satellite laser communication. The optical distortion, spot pattern, and circumferential energy were analyzed and evaluated. The focal-free system was proved to be applicable to a beacon system for tracking and detection.

In 2018, Grajower et al. [75] used a new silicon-based plasma four-quadrant photodetector to operate in the short-wave infrared band; it could accurately track various waist beams at telecommunication wavelengths.

In 2019, Wang et al. [76] proposed a new polynomial-fitting algorithm, using a circular spot with Gaussian energy as the incident-light model. When the radius of the incident spot was 0.5 mm, within a [−0.5 mm, 0.5 mm] detection range, the new fifth-order polynomial-fitting algorithm was adopted, and the maximum error was 1.353 μm. The root-mean-square error was 0.4596 μm, which improved the accuracy of the four-quadrant detector.

In 2020, Liu et al. [77] designed a four-quadrant detector with a five-level amplifier circuit that could be used for low-light detection. The magnification could be adjusted according to the power of the incident light, and the centroid position of the light spot could be intuitively judged. The results showed that the four-quadrant detector had a high resolution, low static noise, and a simple design.

In 2021, Safi et al. [78] studied a surface-to-air free-space optical link for hovering unmanned aerial vehicles (UAVs), and a four-quadrant photodetector array was used at the optical receiver to mitigate the adverse effects of hovering fluctuations by expanding the receiver's field of view.

Table 6.4 shows the development of the four-quadrant photodetector. The above literature introduces the application of the four-quadrant detector to position

Table 6.4 Research progress of four-quadrant detectors

Researcher	Year	Development
Feng	1996	Discussed some of the key factors involved in optical roaming
Kuang	2004	Maximum spot-position detection error was only 0.0277 mm
Wu	2005	Offset of two spot directions was determined
Zhao	2010	Noise of a four-quadrant infrared detector was analyzed
Yang	2013	A suitable optical system was needed to detect the tracking error
Schalkwyk	2014	Influence of the dead zone on detection sensitivity was studied
Tian	2016	Proposed a forward-looking intrinsic day-blind UV four-quadrant detector
Grajower	2018	Proved that a focus-free system can be applied to a beacon system
Wang	2019	Can accurately track all types of beam waists
Liu	2020	A four-quadrant detector with a five-stage amplifier circuit was designed

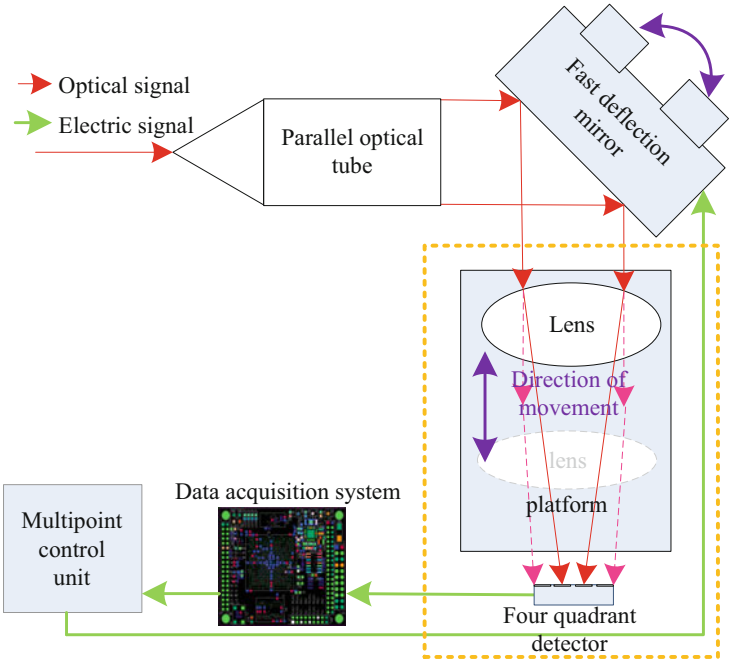


Fig. 6.12 Block diagram of a spot-tracking system

detection and laser collimation, further analyzes the factors that affect the detection accuracy and the causes of related noise, and expounds the improvement in sensitivity of the four-quadrant detector.

Typical Application Scenarios

Figure 6.12 shows a block diagram of a spot-tracking system, based on a four-quadrant detector and fast deflection mirror. The light signal first passes through the

parallel optical tube, and the fast deflection mirror reflects the beacon light from there onto the lens. The lens, with an entry pupil diameter of H, projects the light spot onto the four-quadrant detector used for defocusing.

The data-acquisition system processes the electrical signal and sends it to the multipoint control unit, which controls the deflection angle of the fast deflecting mirror to adjust the incident direction of the optical path, forming a closed-loop tracking system. The mobile electric platform under the lens can control the size of the light spots, and then detect the root mean square error of the infinite integral algorithm under different light spots.

Quantum-Dot Infrared Detectors

Working Principle

The quantum-dot infrared detector is composed of periodic composite layers of quantum dots. Figure 6.13 shows a schematic diagram of the structure of the quantum-dot infrared detector. Each quantum dot in the composite layer has a large number of boundary states, which can receive more electrons. When infrared light is incident on the photosensitive region of the detector, electrons transition from the ground state to the excited or continuous states, which changes the conductivity of the detector, which finally detects the infrared light.

Research Progress

In 1999, Sheng et al. [79] analyzed the noise of a quantum-well infrared detector, including loose-particle noise, thermal noise, generation-compound noise, and 1/f noise, and explored the negative impact of each noise on the detector.

In 2006, Zhao et al. [80] experimentally verified an AlGaAs/GaAs multi-quantum-well infrared detector and showed that the dark current did not appear to change with the temperature when the temperature was less than 40 K. When the

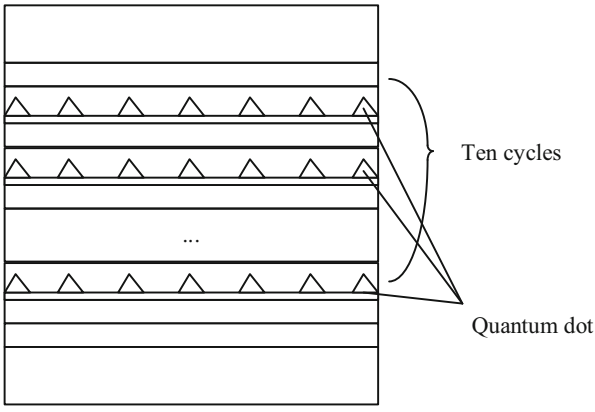


Fig. 6.13 Schematic diagram of the structure of a quantum-dot infrared detector

temperature was higher than 40 K, the dark current increased rapidly with an increase in temperature.

In 2010, Hill et al. [81] proposed a method to extend the detector-cutoff wavelength by combining self-assembled InSb quantum dots with the detector's active region. This extended the cutoff wavelength from 4.2 m to 6 m and exhibited an infrared response at temperatures up to 225 K.

In 2011, Chien [82] developed a model using a quadrant detector at room temperature to optimize a photodetector based on its structure, and the photoelectron capture and transit processes were studied using Monte Carlo simulations.

In 2012, Gustafsson et al. [83] studied quantum dots based on InSb. These quantum dots were used as active materials for a photon detector, which reduced the detector noise.

In 2013, Liu et al. [84] proposed a noise model for a quantum-dot infrared detector. They studied the dark current of the detector, considering the influence of micron-scale and nanoscale electron transports on the excitation energy, and then deduced the noise model.

In 2014, Kim et al. [85] proposed and analyzed a novel detector structure for pixel-level multispectral infrared imaging and then studied the device performance of a raster-integrated quantum-dot well photodetector under back-illumination. By adjusting the grating parameters, a wide range of spectral responses was obtained.

In 2015, Srinivasan et al. [86] used molecular-beam epitaxy to investigate the fabrication of 20-layer silica-doped InAs dots in the structure of quantum-dot infrared photodetector devices.

In 2016, Karow et al. [87] realized a monolithic integrated circuit of a laser and detector, pushed the research on microscale sensor technology to the low-photon quantum limit, and paved the way for on-chip photoelectric-feedback experiments.

In 2017, Kim [88] studied the effects of micron and nanoscale external electrons and electron transport on the noise of quantum-dot infrared detectors by computational simulation. The results showed that under a 2545 kV/cm applied electric field, the noise increased with an increase in the applied electric field and temperature. When the control temperature was less than 80 K, the noise increased very quickly. When the control temperature was greater than 80 K, the noise increased more slowly. Thus, it could be concluded that the lower the temperature, the more obvious the impact of noise on an applied electric field.

In 2018, Wang [89] established a model and algorithm that could represent the dark current of quantum-dot infrared detectors. The model was simplified by combining the actual detector situation. It was found that the temperature, quantum-dot density, and side dimension affected the dark current of the detector.

In 2019, Baira et al. [90] evaluated the size and shape of self-organizing GeSn/Ge quantum dots and found that larger quantum dots with higher aspect ratios might have higher oscillator strengths and longer radiation lifetimes.

In 2020, Gansen et al. [91] investigated a 1/f noise source in a quantum-dot field-effect transistor and produced a 1/f noise model.

Table 6.5 shows some of the progress in quantum-dot infrared detectors. The above studies investigated the influence of the dark current and temperature on noise

Table 6.5 Research progress of quantum-dot infrared detectors

Researcher	Year	Development
Sheng	1999	Loose-particle noise and thermal noise were analyzed
Zhao	2006	Dark current increases rapidly with the increase of temperature
Hill	2010	Extended the detector cutoff wavelength
Chien	2011	Investigated photoelectron capture and transit processes
Lau	2013	Noise model of a quantum-dot infrared detector is proposed
Kim	2014	Wide range of spectral responses was obtained
Karow	2016	Realized a monolithic integrated circuit of laser and detector
Jin	2017	Noise is affected by the applied electric field
Wang	2018	Temperature affects the detector's dark current
Baira	2019	Larger quantum dots with higher aspect ratios
Gansen	2020	Noise model of 1/f noise is derived

generation and the influence of noise on detector performance by analyzing the quantum-dot junctions of quantum-dot infrared detectors.

Bottlenecks and Breakthroughs

Although gallium nitride (GaN), indium antimonide (InSb), mercury cadmium telluride (HgCdTe), and other bulk semiconductor materials can broaden the spectral detection range of CMOS detectors, these bulk semiconductor materials are not compatible with silicon-based CMOS detectors and require flip bonding and other complex processes. This results in a high cost and low pass rate in the detector-manufacturing process. In addition, ultraviolet and infrared photon energies are very different, and cross-band coupling is extremely difficult.

The solution properties of colloidal quantum-dot materials can be used to pattern a quantum-dot thin film after the quantum-dot infrared material is coupled to a CMOS readout circuit. This allows direct lithography to be used in the fabrication of multiband wide-spectrum detectors from ultraviolet to infrared by plane coupling. Using lead sulfide (PbS) quantum dots, the team of Professor Jiang Tang from Huazhong University of Science and Technology has fabricated near-infrared quantum-dot chips and demonstrated good imaging performance under 970 nm infrared irradiation.

To overcome the band bottleneck, Professor Xin Tang from the Beijing Institute of Technology proposed the use of mercury telluride (HgTe) and perovskite colloidal quantum-dot materials to broaden the CMOS detection range, improve the detector-manufacturing scale, realize wafer-level processing, greatly reduce the detector production cost, and improve the process pass rate. Because the bulk material band gap is zero, mercury telluride (HgTe) colloidal quantum dots can be used in short-wave infrared, medium-wave infrared, and long-wave infrared detectors.

This breakthrough greatly simplified the complexity of the optical hardware and software of multiband cameras, reduced the cameras' size and weight, and helped realize the microplatform design of multiband cameras. The team used a direct-photolithography process to form colloid quantum dots into a two-color or

multicolor focal plane array and broke through the traditional block semiconductor-material flip-bonding process to complete multicolor detection and detector coupling. In addition, they realized multiband detection and multicolor imaging capabilities in one detector, greatly reduced the production cost of multicolor detectors, and improved the production-process scale and product pass rate.

Typical Application Scenarios

To better use space resources, countries constantly launch satellites and space stations to establish detection sites, all competing to occupy the optimal positions. Space photoelectric systems play a very important role in this respect, because the new quantum-dot infrared detectors have high detection sensitivity, high working temperature, and antiradiation ability. It would be extremely beneficial to apply quantum-dot infrared detectors to space photoelectric systems.

During laser communications between satellite and ground or satellite and satellite, the space environment is complicated because of the long transmission distance, which mainly causes two problems: The first is that the intensity attenuation intensifies as the communication distance increases. The second is the harsh space environment, which requires photoelectric devices to have better temperature characteristics and radiation-resistance performance. A quantum-dot infrared detector not only has a high working temperature and radiation resistance but also has advantages in structure and volume. Therefore, it has broad application prospects in space environments.

Balanced Detectors

Working Principle

Figure 6.14 shows the principle diagram of a balanced detector receiving coherent light [92]. When the local light and signal light fully meet the coherence conditions, the light field interferes and the photodetector converts the light to an electrical

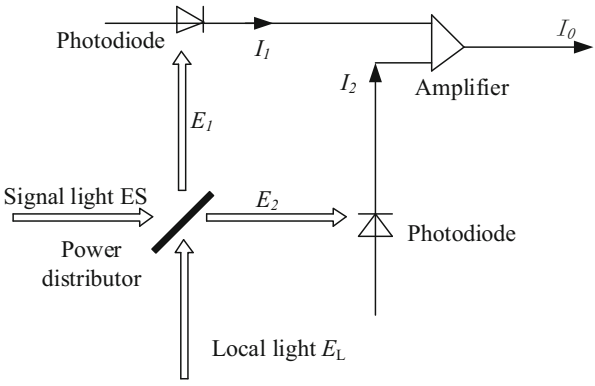


Fig. 6.14 Block diagram of a coherent-light-receiving balanced detector [92]

signal. The signal light is represented by E_S, and the local oscillator light is represented by E_L. Assume that the power divider is an ideal device, that is, linear, symmetrical, and lossless:

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \exp(j\psi_r) \begin{bmatrix} 1 & \exp\left(j\frac{\pi}{2}\right) \\ \exp\left(j\frac{\pi}{2}\right) & 1 \end{bmatrix} \begin{bmatrix} E_S \\ E_L \end{bmatrix}, \tag{6.13}$$

where w_{IF} denotes the reflected phase shift of the power divider, E_S denotes the signal light, and E_L denotes the local oscillator light.

$$I_1(\mathrm{t}) - I_2(\mathrm{t}) = 2\frac{\eta e}{h\nu}\sqrt{P_S P_L}\cos(w_{IF}t + \phi_S - \phi_L), \tag{6.14}$$

where w_{IF} is the frequency of the intermediate frequency signal; ϕ_S and ϕ_L are the phases of the signal light and local oscillator light, respectively; P_S and P_L are the optical powers of the signal light and local oscillator light, respectively; η is the quantum efficiency of the photodetector; e is the amount of charge; h is Planck's constant; ν is the optical frequency; and I_0 is amplified by the amplifier to generate the current. Because the balanced detector can suppress the common-mode signal, the noise is reduced, and the detector sensitivity is improved.

Research Progress

In 2014, Dai et al. [93] designed a set of high-speed balanced-detection systems that achieved coherent detection at different rates. For a 2.5-Gbit/s signal rate, the direct-detection sensitivity reached −26 dBm and the dynamic range was −26–3 dBm. When the local oscillator light was 3 dBm, the sensitivity reached −46.5 dBm and the dynamic range was −46.5–3 dBm.

In 2015, Guo et al. [94] analyzed the output noise of balanced heterodyne detection and established a relevant noise model. The simulation results showed that, compared with the signal-to-noise ratio of common coherent detection, the signal-to-noise ratio of a balanced heterodyne-detection method could be improved by reasonably setting the structural parameters of the detector.

In 2016, Shi et al. [95] designed a coherent-detection test scheme based on the composition of a balanced detector. When the communication rate was 5 Gbit/s, the common-mode rejection ratio was higher than 27 dB. Balanced detection was adopted to achieve coherent demodulation, and the minimum signal optical power was improved by 8 dB compared with single-tube detection.

In 2017, Munshid [96] designed and constructed a heterodyne detector based on photonic crystal fiber and found that the relationship between the wavelength shift and temperature change was linear.

In 2018, Yong et al. [97] estimated the dispersion of coherent optical communication systems based on heterodyne detection and investigated the nonlinear limitations of 10-Gbit/s nonreturn to zero electron predistortion systems by simulation.

In 2019, Zhang et al. [98] designed a balanced detector in a high-speed terahertz time-domain spectroscopy system. Two low-noise photodiodes in series were used

to directly detect the difference in the photocurrents of two laser beams and reduce the detector noise.

In 2021, Hu et al. [99] studied a novel heterodyne detection scheme and compared it to the heterodyne detection of a single-ended photodetector. The new scheme allowed for the use of a simple transmitter structure, reducing the receiver bandwidth to a value far below the sign rate, while preserving the ability to compensate for digital dispersion.

As shown in Table 6.6, the aforementioned literature focuses on the application of balanced detectors in coherent optical communication systems based on heterodyne detection. The structure can suppress the noise, and the noise model of the detectors is obtained simultaneously.

Typical Application Scenarios

The principle of coherent detection is illustrated in Fig. 6.15. The light emitted from the laser acts as a carrier and is modulated by the signal either directly or externally. At the receiving end, the spatial optical signal W_s is coupled into the fiber via a fiber coupler, and the lightwave generated by the local oscillator W_L is superimposed on the received signal in the mixer, producing an intermediate frequency (IF) signal at the output end of the balanced detector. Zero-difference detection can be directly obtained from the baseband signal and balanced-detector output signal after IF amplification, filtering, demodulation, and restoration of the digital signal to the sending end.

Table 6.6 Research progress of balanced detectors

Researcher	Year	Development
Bae	1993	Excess intensity noise of local vibration light is suppressed
Dai	2014	Sensitivity can reach −46.5 dBm
Guo	2015	Improved SNR ratio of the balanced heterodyne detection method
Shi	2016	Common-mode rejection ratio is higher than 27 dB
Munshid	2017	Relationship between wavelength shift and temperature is linear
Yong	2018	Ten Gbit/s non-return-zero electron predistortion systems
Zhang	2019	Detector noise is reduced
Hu	2021	Receiver bandwidth is reduced to a value well below the sign rate

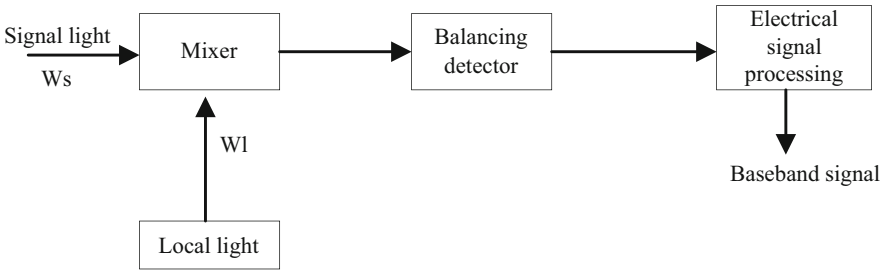


Fig. 6.15 Schematic diagram of coherent detection

Double Balanced Probes

Working Principle

Double-balanced detection is based on the structural composition of a balanced heterodyne detection system. The double-channel balanced receiving technology is shown in Fig. 6.16, which uses four photodetectors that are almost identical. The two detectors at the receiving end are a group, forming two groups of balanced heterodyne detectors [100].

The signal light and local light enter the port at the input end of a 90° optical mixer. After a quarter of the waveplate, the local light can be converted into circularly polarized light. At this time, the two present a 45° angle. Through the beam-splitting prism, the signal and local lights can be shot into half of the wave plate after the mixed-beam splitting; at this time, the two will present a 22.5° angle. Finally, four light beams are obtained, based on the polarization beam splitter. Two orthogonal IF current signals can be obtained by combining two beams of light with opposite phases and inputting them into a balanced detector.

$$\begin{cases} I : I_0 - I_{180} = 2k_1k_4 \cos\left[\omega_{IF} - \left(\varphi + \frac{\pi}{4}\right)\right] \\ Q : I_{90} - I_{270} = 2k_2k_3 \sin\left[\omega_{IF} - \left(\varphi + \frac{\pi}{4}\right)\right] \end{cases}, \tag{6.15}$$

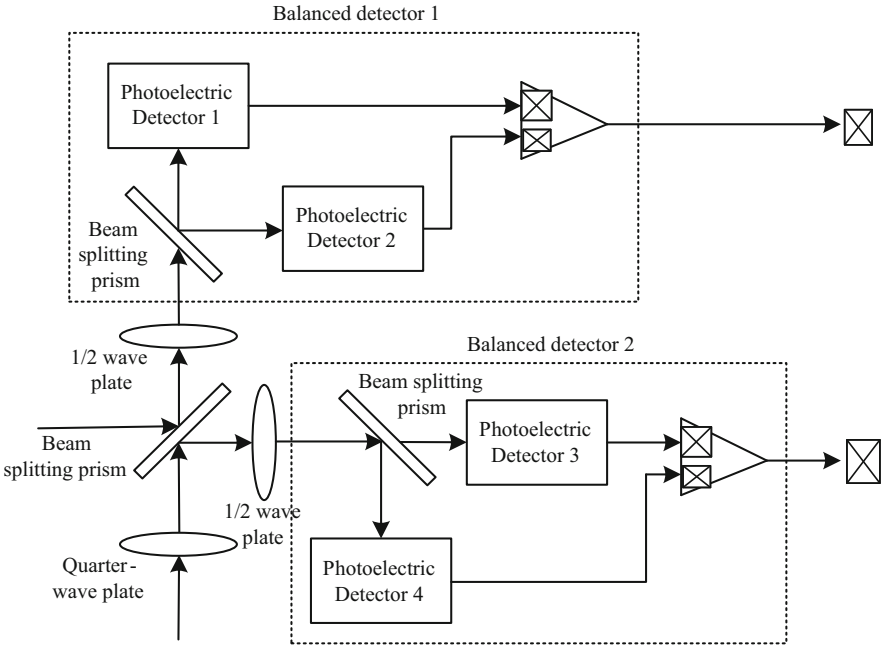


Fig. 6.16 Double-balanced probe

where I and Q are the 90° phase difference between the two channels, that is, the in-phase component and the orthogonal component required by the demodulation information. I_0, I_{180}, I_{90}, and I_{270} are the four optical-mixer outputs obtained through the polarization splitter prism; k_1, k_2, k_3, and k_4 represent the local light and signal light in the directions of the component, φ is the phase difference, and ω_{IF} represents the intermediate frequency. When the two subtracted photocurrents after the DC component are completely offset, and the intensity noise related to the DC component is also eliminated, the AC term is proportional to the square root of the local light power. Therefore, the noise intensity has a much smaller influence.

Research Progress

In 2003, Li et al. [101] verified the physical mechanism of the balanced detector and built a 1550 nm balanced heterodyne detection system. The results showed that the signal-to-noise ratio performance of the double-balanced detector was much better than that of the single-ended detector.

In 2010, Gao et al. [102] applied dual-balanced detection technology to a weak signal-detection field, for example, in Lidar, the polarization factor of incident light will cause a polarization mixing error, and the error of the dual-balanced detection system was discussed. The results showed that the vertical orthogonal factor of the ellipse formed by polarized light on the polarization splitter was 0.523, when the polarization degree was in the range of (0, 0.02).

In 2013, Zhu [103] established a dual optical-path-balanced detection system using 3 × 3 couplers to solve the corrosion and electromagnetic-interference problems of traditional acoustic microphones. This enabled the differential processing of two equal-amplitude signals. Experiments showed that the system was stable and the waveform distortion was small. It extracted and recovered speech signals in the 300–3400 Hz range.

In 2013, Huan [104] studied the weak laser-signal detection of a double-balanced heterodyne detection system, starting from the aspects of the system structure and a mathematical model. Through simulation analysis, they deduced a conclusion that conformed to the mathematical model, which laid the foundation for research on weak laser-detection technology.

In 2015, Zheng et al. [105] proposed a new dual-balanced differential configuration method with a photon-detection efficiency of 10% for a 1550 nm laser at 10 MHz. The results showed that this configuration effectively resolved ultraweak avalanche signals.

In 2015, Peng et al. [106] introduced a new photoelectric heterodyne-receiving technology based on double-balanced coherent detection, explained the working principle of double-balanced coherent detection, and further analyzed double-balanced and balanced coherent-detection systems.

In 2016, Li et al. [107] designed a double-balanced sound-receiving system based on the Mach–Zehnder ring structure to increase the measuring distance of a laser Doppler velocimeter. The SNR expression of the double-balanced sounding system was derived. Simulation results showed that the signal-to-noise ratio of the double-balanced detection system was higher than that of common coherent detection.

In 2017, Huang et al. [108] considered the shortcomings of common ground-fault detection methods and proposed a new two-bridge detection method. This method can be used for fault detection, which not only solves the problem caused by the dead zone but also does not affect the distributed capacitance. The feasibility of this method was experimentally verified.

Table 6.7 shows the development of the dual-balanced detector. The above literature describes the working principle of double-balanced detection technology and analyzes whether it can be applied to the detection of weak signals and other fields. A comparison with common coherent detection technology shows that double-balanced detection technology can improve the system signal-to-noise ratio.

Typical Application Scenarios

Figure 6.17 presents a simulation block diagram of a coherent-optical communication system. In the simulation, a pseudorandom bit sequence forms the modulated signal, and the signal light is modulated by quaternary phase-shift keying (QPSK). In the experiment, the oscillator light and signal light are introduced into the mixer for mixing. The mixer signal is output to the double-balanced detection system, and its signal is converted to analog after digital signal processing. The receiver decision circuit measures the signal-to-noise ratio to determine whether the system performance is good or bad by observing the bit-error rate.

Table 6.7 Research progress on double-balanced detectors

Researcher	Year	Development
Li	2003	Improved the SNR performance of double-balanced detectors
Gao	2010	Applied dual-balanced technology to weak signal detection
Zhu	2013	Used a 3 × 3 coupler to establish a balanced detection system
Huan	2013	Weak laser-signal detection in a double-balanced detection system
Zheng	2015	Proposed a double-balanced differential configuration method
Peng	2015	Studied double-balanced coherent detection
Li	2016	Created a DB detection system based on the Mach–Zehnder ring structure
Huang	2017	Proposed a new two-bridge detection method

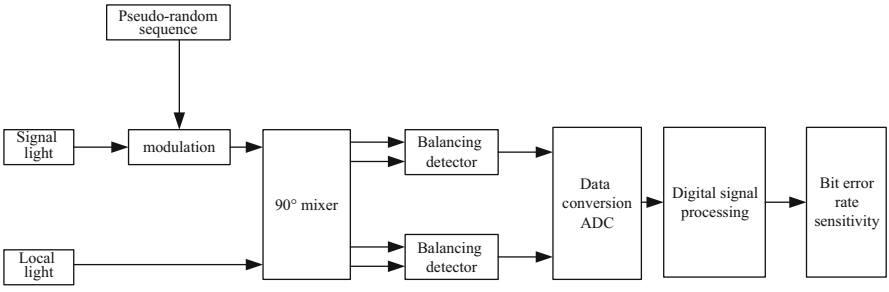


Fig. 6.17 Schematic diagram of a coherent-optical communication-system simulation

Photodetector Noise

Sources of Photodetector Noise

Loose-Particle Noise

When incident light irradiates a detector surface, the photoelectrons oscillate randomly, and scattered noise is generated. The electrons emitted by the electron tube in any short time τ will fluctuate near the mean value. From the perspective of the mean-square deviation of the fluctuation, the power of the scattered particle noise is

$$i_n^2 = 2eI\Delta f, \tag{6.16}$$

where i_n^2 represents the power of the scattered particle noise, e represents the electron charge, Δf represents the noise bandwidth, and I is the average photocurrent generated by the detector.

For the cases of incident light and no incident light, the dark-current noise power and light-radiated scattered-particle noise power are, respectively,

$$i_{nd}^2 = 2eI_d\Delta f \tag{6.17}$$

$$i_{np}^2 = 2eI_p\Delta f, \tag{6.18}$$

where i_{nd}^2 denotes the dark-current noise power, i_{np}^2 denotes the light-radiated scattered-particle noise power, e denotes the amount of charge in an electron, and Δf represents the noise bandwidth of the detector. The average photocurrent without incident light is represented by I_d, and the average photocurrent with incident light is represented by I_p.

Generation-Compound Noise

The constant movement of charge carriers in a semiconductor leads to the random generation and recombination of electrons and holes; hence, the average carrier concentration will fluctuate continuously. The noise generated at this time is called generation-compound noise, and the expression is

$$i_{ng-r}^2 = 4eIM\Delta f, \tag{6.19}$$

where the power generating the generation-compound noise is represented by i_{ng-r}^2, I represents the average photocurrent, Δf represents the noise bandwidth of the detector, and M represents the gain.

Thermal Noise

The random motion of the charge carriers in the dissipative element produces thermal noise. For resistances, these electrons produce thermal noise as long as the resistance is in thermal equilibrium, even if no voltage is applied.

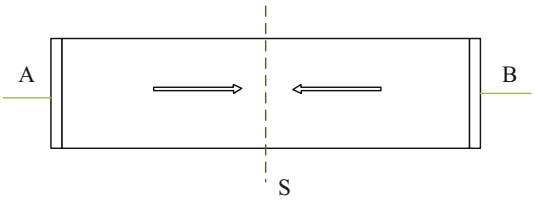


Fig. 6.18 Equivalent diagram of an AB electrode

An AB-electrode equivalent diagram is shown in Fig. 6.18. The resistance between the poles of AB is denoted by R. At an absolute temperature T, the electrons between the poles of AB are in constant random thermal motion.

Because there are voltage fluctuations at both ends of AB, the average power of the thermal noise obtained is

$$U_{nJ}^2 = 4KT\Delta fR, \tag{6.20}$$

where U_{nJ}^2 denotes the average power of thermal noise, the Boltzmann constant is denoted by k, the measurement bandwidth is denoted by Δf, the absolute temperature is denoted by T, and R is the resistance between the AB electrodes. If the noise is caused by the current,

$$i_{nJ}^2 = \frac{4KT\Delta f}{R}, \tag{6.21}$$

where i_{nJ}^2 represents the average power of the thermal noise.

The root-mean-square values of the thermal-noise current and voltage are typically calculated as follows:

$$\sqrt{i_{nJ}^2} = \left(\frac{4KT\Delta f}{R}\right)^{1/2} \tag{6.22}$$

$$\sqrt{U_{nJ}^2} = (4KT\Delta fR)^{1/2}, \tag{6.23}$$

where $\sqrt{i_{nJ}^2}$ is the root mean square value of the thermal noise current and $\sqrt{U_{nJ}^2}$ is the root mean square value of the thermal noise voltage.

1/f Noise

The detector-manufacturing process introduces unnecessary trace impurities or small nonuniform particles into the sensitive layer. Under the action of current, microburst pulses are generated between these particles, resulting in 1/f noise.

The current mean-square value of 1/f noise can be expressed as

$$i_{nf}^2 = k_1 \frac{I^b \Delta f}{f^a}, \tag{6.24}$$

where i_{nf}^2 represents the current mean square value of 1/f noise, and the proportionality coefficient is k_1. k_1 is affected by various factors, including the detector manufacturing process, electrode contact, and size of the detector. Δf is the working bandwidth, a is a constant related to the material of the detector, usually between 0.8 and 1.3. Current I is in the device; b is related to I, b usually taking the value 2.

Typical Photodetector-Noise Models

Noise Model of an Avalanche Photodetector

Thermal Noise

For an avalanche photodiode, the thermal noise current satisfies Eq. (6.18), if the resistance inside the diode is ignored, and only the effect of the load of the current in which the diode is located is considered.

Loose-Particle Noise Generated by the Dark Current

The continuous random movement of carriers in the P-N junction generates loose particle noise. Then [109]

$$i = i_p + i_s, \tag{6.25}$$

where i is the total current flowing through the P-N junction. The scattered-particle noise current is represented by i_s and the average value of the current is i_p. If the bandwidth of the receiving system is B, the effective value of the scattered noise can be expressed as

$$i_s = \sqrt{2ei_pB}, \tag{6.26}$$

where i_s is the effective value of the scattered particle noise, e is the electronic charge, and the average current value is i_p. In a dark environment, for an avalanche photodiode with a multiplication factor M, the dark current is

$$i_d = i_b + Mi_t, \tag{6.27}$$

where i_d is the dark current, i_b is the surface leakage current, i_t is the volume leakage current, and M is the multiplication factor. The dark current of the detector causes loose-particle noise, and the effective value of the noise current is expressed as [109]

$$i_s' = \sqrt{2ei_tBM^2 + 2ei_bB}, \tag{6.28}$$

where i_s' is the effective value of the noise current.

Excess Noise

The gain of an avalanche photodiode represents the ratio of the output photocurrent to the incoming photocurrent [109]:

$$M = \frac{I_p - I_d}{I_{p0} - I_{d0}}, \tag{6.29}$$

where the multiplication factor is M, I_p is the photocurrent of the avalanche photodiode, I_d is the dark current of the avalanche photodiode, I_{p0} represents the photocurrent at unit gain, and I_{d0} represents the dark current at unit gain.

Assuming that the avalanche multiplication region is uniform, the overnoise factor is related to the gain and ionization coefficient, that is,

$$F(M) = kM + (1-k)\left(1 - \frac{1}{M}\right), \tag{6.30}$$

where $F(M)$ is the overnoise factor, k is the ionization coefficient, and M is the multiplication factor. In device design, materials with smaller k values should be selected for the multiplication region as much as possible.

PIN Photodetectors

Thermal Noise

Thermal noise is the current or voltage fluctuation caused by the random movement of charge carriers in the material at a certain temperature. The thermal noise of a PIN detector satisfies Eq. (6.20).

Loose-Particle Noise

The expression for the scatter-noise current satisfies Eq. (6.14).

Generation-Compound Noise

The mean-square current of PIN-detector generation-compound noise is

$$i_{nf}^2 = \frac{cI^\alpha}{f^\beta}\Delta f, \tag{6.31}$$

where i_{nf}^2 is the mean square value of the current of generation-compound noise, the average value of the output current is expressed by I, the operating frequency is expressed by f, α is approximately 2, Δf is the operating bandwidth, and c is the proportionality constant. The values of α, β, and c can be measured empirically and are related to variations in the device surface.

1/f Noise

The mean-square current of the 1/f noise of a PIN detector is expressed as

$$i_{nf}^2 = \frac{cI^\alpha}{f^\beta}\Delta f, \tag{6.32}$$

where i_{nf}^2 is the mean square value of the current of 1/f noise, the working frequency is expressed by f, α is approximately 2, between 0.8 and 1.5., Δf is the operating bandwidth, c is a constant of proportionality, and the values of a α, b β, and c can be measured experimentally and are related to the surface undulation state of the device.

Photomultiplier Tubes

Signal-to-Noise Ratio

Suppose the photomultiplier is order K, and the average number of electrons collected by its anode is n_a. If each order of multipliers is the same and conforms to the Poisson distribution, [110]

$$\mathrm{SNR}_a \approx \left|\frac{\eta n_p(\delta-1)}{\delta}\right|^{1/2}, \tag{6.33}$$

where SNR_a is the signal-to-noise ratio of the photomultiplier tube, ηn_p is the average number of photoelectrons, and δ is the gain. If $\delta \gg 1$, then

$$\mathrm{SNR}_a \approx \sqrt{\eta n_p}. \tag{6.34}$$

Scattered-Particle Noise Current

The mean square current value of the scattered noise is [110]

$$I_{Npe}^2 = 2q^2\eta I_p \Delta f, \tag{6.35}$$

where the charge is q and forms a series of current pulses. The spectrum of the current pulse consists of many harmonics; Δf is the harmonic interval, I_p is the rate reached by the average photon, and η is the quantum efficiency. The mean-square current of the scattered-particle noise of the photomultiplier tube is

$$I_{N_a} = 2qI_{pe}\Delta f\left(1+\frac{1}{\delta 1}+\frac{1}{\delta 2}+\cdots+\frac{1}{\delta_1\delta_2\cdots\delta_k}\right)m_k, \tag{6.36}$$

where $m_k = \prod_{i=1}^{k}\delta_i$ is the total gain of the photomultiplier and I_{pe} is the photocathode current.

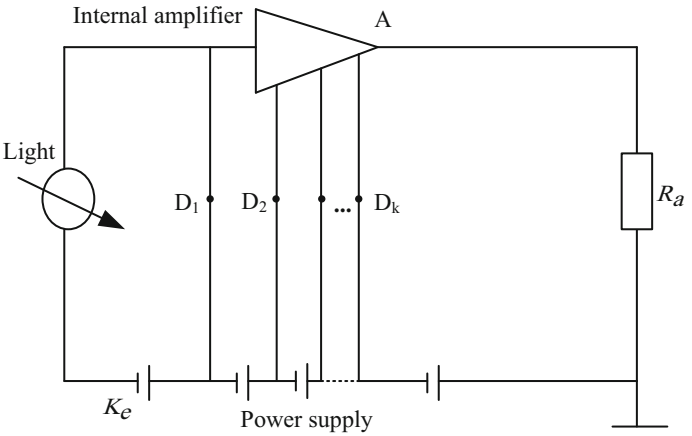


Fig. 6.19 Equivalent-circuit diagram of a photomultiplier

Simplified Photomultiplier Model

The simplified noise model of a photomultiplier consists of a photocathode, several secondary emission multipliers, and an anode. An equivalent circuit diagram of the photomultiplier is shown in Fig. 6.19.

In Fig. 6.19, D_1, D_2, ... D_k represent the first k multipliers, K_e is the photocathode, A is the output anode of the inner amplifier, and R_a is the anode load resistor.

The photomultiplier tube can also be viewed as a current source whose anode current is an m_k multiple of the photocathode current:

$$I_a = m_k I_{pe} = \left(\prod_{i=1}^{k} \delta_i \right) I_{pe}, \tag{6.37}$$

where I_{pe} is the cathode current and $\prod_{i=1}^{k} \delta_i$ is the total gain of the photomultiplier tube.

In Fig. 6.20 [110], R_a is the anode resistance, I_{pe} is the cathode current, $m_k = \prod_{i=1}^{k} \delta_i$ is the total gain of the photomultiplier, I_p is the average photon-arrival rate, C_a is the anode capacitance, R_p is the cathode resistance, C_p is the cathode capacitance, and U_a is the anode voltage source.

Four-Quadrant Detector

Thermal Noise

The thermal noise of the four-quadrant detector satisfies Eq. (6.20).

1/f Noise

The 1/f noise-voltage power-spectrum density in the four-quadrant detector is expressed as follows [111]:

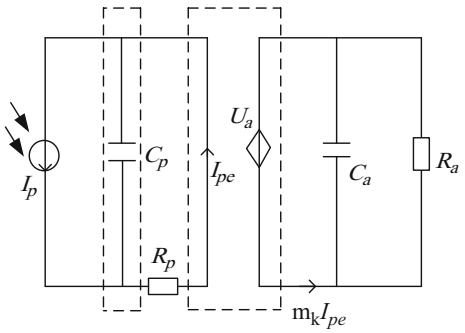


Fig. 6.20 Simplified photomultiplier noise model

$$S_V(f) = AV^{\beta}/f^{\gamma}, \tag{6.38}$$

where $S_V(f)$ is the power spectrum density of the 1/f noise voltage, A is related to the detector material and is a constant, β is the exponential factor, V is the noise voltage, f is the frequency, and γ is the frequency exponent.

In 1969, Hooge proposed his famous empirical formula and mobility fluctuation model:

$$\frac{S_I(f)}{I^2} = \frac{S_R(f)}{R^2} = \frac{S_v(f)}{V^2} = \frac{\alpha_H}{fN}, \tag{6.39}$$

where the power spectral density of the noise current is represented by $S_I(f)$, the power spectral density of the resistance is represented by $S_R(f)$, and the power spectral density of the voltage is represented by $S_v(f)$. The Hooge coefficient is α_H; it can be used to measure the 1/f device noise level, where f is the frequency and N is the total number of carriers.

Generation-Compound Noise

When generating four-quadrant detector generation-compound noise, the power spectral density is satisfied as follows [111]:

$$S_I(f) = \sum_{i=1}^{m} \frac{C_i \tau_i}{1 + \left(2\pi f_{\tau_i}\right)^2}, \tag{6.40}$$

where $S_I(f)$ is the power spectral density of generation-compound noise, τ_i is the characteristic time constant of generation-compound noise, m is the number of deep energy levels, C_i represents the amplitude of the generation-compound noise component of time constant τ_i, and f_{τ_i} is the frequency of the characteristic time constant.

The characteristic time constant τ_i can be expressed as

$$\frac{1}{\tau_i} = CT^2 \exp(-Ea/kT), \tag{6.41}$$

where k is Boltzmann's constant, C is a constant, Ea is the activation energy of the deep-level impurity, and T is the temperature.

Quantum-Dot Detector

Dark Current

The barrier has moving carriers, and the dark current can be calculated from the density of the carriers [84]:

$$I_{\text{dark}} = 2evuEA\left(1+\left(\frac{uE}{v_s}\right)^2\right)^{-1/2}\left(\frac{m_b kT}{2\pi h^2}\right)^{3/2}\exp\left(-\frac{E_{0,\text{micro}}\exp(-E/E_0)+E_{0,\text{nano}}-\beta E}{KT}\right) \tag{6.42}$$

where e is the electronic charge, v is the electron drift velocity, A is the area of the detector, E is the bias electric-field intensity and can be approximated by dividing the applied voltage by the thickness of the intrinsic region of the detector, μ is the electron mobility, v_s is the maximum moving velocity of an electron, m_b is the effective mass of an electron, k is Boltzmann's constant, T is the temperature, and h is a normalized Planck's constant. $E_{0,\text{micro}}$ is the excitation energy of a micron-scale electron transport under zero pressure bias, $E_{0,\text{nano}}$ is the excitation energy of a nanoscale electron transport under zero pressure bias, E_0 describes how fast or slow the excitation energy of the micron-scale electron transport varies with the electric field, and β describes how fast or slow the excitation energy of the nanoscale electron transport varies with the electric field.

Noise Gain

The noise gain can be calculated by the ratio of the electron recombination time τ_r to the electron transit time τ_d [84]:

$$g_n = \frac{\tau_r}{\tau_d}, \tag{6.43}$$

where g_n is the noise gain.

In general, the electron capture by quantum dots is considered to be symmetric, and the specific calculation form is as follows:

$$\frac{1}{\tau_T} = N_t(4\pi D R_t)\left\{1-\left[\frac{D}{V_t R_t^2}\tan h\left(\frac{V_t R_t^2}{D}\right)\right]^{1/2}\right\}, \tag{6.44}$$

where R_t is the effective radius of the quantum dot, V_t is the capture speed of a quantum dot capturing an electron, τ_T is the capture time of a quantum dot capturing an electron, N_t is the quantum number of points, and D is the band-diffusion coefficient, which is related to electron mobility μ:

$$\mu = \frac{eD}{kT}. \tag{6.45}$$

Equations (6.40) and (6.41) are substituted into Eq. (6.39) and sorted to obtain the noise gain, which can be written as

$$g_n = \frac{uE\left(1 + \left(\frac{uE}{v_s}\right)^2\right)^{-1/2}}{N_t(4\pi R_t ukT/e)\left\{1 - \left[\frac{ukT}{ev_t R_t^2}\tan h\left(\frac{ev_t R_t^2}{ukT}\right)\right]^{1/2}\right\}}, \tag{6.46}$$

where E is the bias electric-field intensity of the quantum-dot infrared detector, μ is the mobility of the electron, and v_s is the saturation movement speed of the electron.

Noise

The noise of quantum-dot detectors mainly comes from the electron production–recombination process and can be written as

$$i_n = \sqrt{4eg_n I_d} \tag{6.47}$$

where e is the basic electron charge and g_n is the noise gain, which is approximately equal to the photoconductance gain, and I_d is the dark current.

According to the dark-current and gain-calculation methods, Eqs. (6.39) and (6.43) are substituted into Eq. (6.44) to obtain the formula for calculating quantum-dot infrared detector noise [84]:

$$i_n = \sqrt{\frac{8e^2u^2E^2\left(1 + \left(\frac{uE}{v_s}\right)^2\right)^{-1}\left(\frac{m_b kT}{2\pi h^2}\right)^{3/2}\exp\left(-\frac{E_{0,\text{micro}}\exp(-E/E_0) + E_{0,\text{nano}} - \beta E}{KT}\right)}{N_t(4\pi R_t ukT/e)\left\{1 - \left[\frac{ukT}{ev_t R_t^2}\tan h\left(\frac{ev_t R_t^2}{\mu KT}\right)\right]^{1/2}\right\}}} \tag{6.48}$$

where the bias electric-field intensity is expressed by E, m_b is the effective mass of the electron, μ is the electron mobility under a zero voltage bias, $E_{0,\text{micro}}$ is the excitation energy generated with a micron-magnitude electron transmission, and $E_{0,\text{nano}}$ is the excitation energy generated with a nanometer-magnitude electron transmission. The size will be affected by the change of the electric field, β describes the change of the speed, and v_s represents the maximum speed that can be reached when the electron is moving.

Balanced Detector

Thermal Noise

The thermal noise of the balanced detector satisfies Eq. (6.20).

Loose-Particle Noise

The mean-square value of the scattered noise current is expressed as

$$i_{sn}{}^2 = 2eP_LB, \tag{6.49}$$

where e denotes the electron charge and B denotes the noise bandwidth.

Double-Balanced Detector

Loose-Particle Noise

The power spectral density of a single sideband of scattered noise is expressed as [112]

$$G_{\text{shot}}^2 = 2 \cdot q \cdot (\eta \cdot P + I_d) \cdot \Delta f, \tag{6.50}$$

where G_{shot}^2 is the single-sideband noise power spectral density, q is the electron charge, η is the device-response factor, I_d corresponds to the dark current, Δf is the single-sideband effective noise bandwidth, and P is the total received power.

Thermal Noise

The main source of thermal noise is cross-resistance amplification; the noise can be represented as follows [112]:

$$\sigma_{\text{thermal}}^2 = (i_{\text{TIA}})^2 \cdot \Delta f, \tag{6.51}$$

where $\sigma_{\text{thermal}}^2$ represents thermal noise, i_{TIA} is the single-side differential input noise current either, and Δf is the single-sideband effective noise bandwidth.

Relative-Intensity Noise

The relative-intensity noise caused by local vibration light can be expressed as [112]

$$\sigma_{\text{RIN}} = 2(\eta P_L)^2 \cdot \text{RIN} \cdot \Delta f, \tag{6.52}$$

where σ_{RIN} represents the relative intensity noise, P_L is the local vibration power, RIN is the relative-intensity noise coefficient, and η is the equipment-response coefficient.

Summary

In this chapter, specific detectors were analyzed, including the avalanche photodetector, PIN photodetector, photomultiplier, four-quadrant detector, quantum-dot infrared detector, and balanced detector. Their thermal noise, loose-particle noise, generation-compound noise, and 1/f noise were analyzed from the noise-generation mechanism, and the noise model of related detectors was established. This provided

a basis for studying the signal quality at the receiving end of an optical wireless communication system.

In the future, the following work is worthy of further study and discussion.

1. Develop new composite materials with a high carrier mobility, longer lifetime, and diffusion distance, as well as a large photoconductance gain, to improve the response speed and quantum efficiency of the detector.
2. Explore new integration methods to increase the photodetector integration and reduce the number of detectors and search time. In addition, the resolution of the detector-channel wavelength should be enhanced and the cost reduced.
3. Further reduce the carrier concentration, noise, and impact of the junction area and pressure-welding size on the dark current of the detector, and thereby improve the sensitivity of the photodetector.
4. Improve the detector-surface process state, avoid the existence of uneven particles or unnecessary trace impurities in the photosensitive layer, and reduce the 1/f noise of the detector.
5. Improve the detector structure, reduce the composite loss of photogenerated carriers in the dead zone of the device surface, and improve the detector detection efficiency.
6. Ultrahigh speed, low-voltage, low-loss lithium-niobate thin-film electro-optic modulators have introduced a new era in the sustainable and rapid development of information technology. Whether future optical-interconnection technology can transmit the massive amounts of information in sufficient time is a fundamental issue concerning the sustainability of information-technology development.

References

1. Ke X Z, Dong K, Optical-Wireless Communication [M], Beijing: Springer Press, 2022: 10–13.
2. Pauwels H. Phase and amplitude fluctuations of the laser oscillator [J]. Quantum Electronics IEEE Journal of, 1966, 2(3):54–62.
3. Leeson D B. A simple model of feedback oscillator noise [J]. Proc of IEEE, 1966, 54(2): 329–330.
4. Van D Z A. Noise in solid-state devices and lasers [J]. Proceedings of the IEEE, 1970, 58(8): 1178–1206.
5. Harder C, Katz J, Margalit S, et al. Noise equivalent circuit of a semiconductor laser diode [J]. IEEE Journal of Quantum Electronics, 1982, 18(3):333–337.
6. Petermann K. Laser Diode Modulation and Noise [M]. Sold and distributed in the U.S.A. and Canada by Kluwer Academic Publishers, 1991: 56–60.
7. Kuo C Y. Fundamental second-order nonlinear distortions in analog AM CATV transport systems based on single frequency semiconductor lasers [J]. Journal of Lightwave Technology, 1992, 10(2):235–243.
8. Bihan J L, Yabre G. FM and IM intermodulation distortions in directly modulated single-mode semiconductor lasers [J]. IEEE Journal of Quantum Electronics, 1994, 30(4):899–904.
9. Chai Y J, Yang Z X. Research on Influence of Laser Phase Noise on Characteristics of high-speed IM/DD Fiber Communication System [J]. Journal of Communications, 1997, 18(7):1–5.
10. Fukuda M, Hirono T, Kurosaki T, et al. 1/f noise behavior in semiconductor laser degradation [J]. IEEE Photonics Technology Letters, 2002, 5(10):1165–1167.

11. Lawrence J S, Kane D M. Nonlinear dynamics of a laser diode with optical feedback systems subject to modulation [J]. IEEE Journal of Quantum Electronics, 2002, 38(2):185–192.
12. Yu B L, Zhen S L, Zhu J, et al. Experimental study on low noise fiber laser [J]. Acta Optica Sinica, 2006, 26(2):217–220.
13. Domenico G D, Stéphane Schilt, Thomann P. Simple approach to the relation between laser frequency noise and laser line shape [J]. Applied Optics, 2010, 49(25):4801–4807.
14. Liu J H. Influence of Laser Phase Noise on Performance of Coherent Fiber communication System [J]. Journal of Xi'an University of Posts and Telecommunications, 2012, 17(5):25–28.
15. Mahmoud S W Z, Mahmoud A, Ahmed M. Noise performance and nonlinear distortion of semiconductor laser under two-tone modulation for use in analog CATV systems [J]. International Journal of Numerical Modelling Electronic Networks Devices & Fields, 2015, 29(2):280–290.
16. Chen D, Ke X Z, Zhang L. Intermodulation distortion characteristics of laser in turbulent channel [J]. Acta Photonica Sinica, 2016, 45 (2): 99–103.
17. Chen J N. Experimental Study on Key Technologies of Subcarrier Modulation Non-Optical Domain Extrinsic Error Detection in Wireless Optical Communication [D]. Xi'an University of Technology, 2016.
18. Song Z Y, Yao G B, Zhang L L, et al. Influence factors of phase noise in single frequency fiber laser [J]. Infrared and Laser Engineering, 2017, 46(3):90–93.
19. Bai Y. Characterization method of laser linewidth at 2 μm band and development and application of single longitudinal mode thulium doped fiber laser [D]. Beijing Jiaotong University, 2019.
20. Du Y C, Zhang R, Wang L S, et al. Chaotic Synchronization of Distributed Bragg Reflection Semiconductor Laser Driven by Common Noise [J]. Acta Optica Sinica, 2022, 42(23): 2314003.
21. Shun Y. Photomultiplier tubes for laser ranging [J]. Laser and Infrared, 1972, 2(8):35–37.
22. Li G. Polarization Effects of Grating Monochromator and Photomultiplier Tubes [J]. Journal of Qufu Normal University. 1989, 2(1):34–40.
23. Zhou R F. Outlook of photomultiplier tubes [J]. Optoelectronics Technology, 2000, 4(2): 84–89.
24. Guo C L, Zeng D, Li J, et al. Noise modeling of photomultiplier tubes [J]. Nuclear Electronics and Detection Technology, 2004, 2(1):117–120.
25. Chowdhuri M B, Ghosh J, Manchanda R, et al. Measurement of spatial and temporal behavior of H(α) emission from Aditya tokamak using a diagnostic based on a photomultiplier tube array [J]. Review of Scientific Instruments, 2014, 85(11): 101–108.
26. Maltseva Y, Emanov F A, Petrenko A V, et al. Distributed beam loss monitor based on the Cherenkov effect in an optical fiber [J]. Physics-Uspekhi 2015, 58(5): 516–519.
27. Chen X Q, Tang Q Y, Wang X H. Development of Small Radiation Measuring Probe Based on Silicon Photomultiplier Tube [J]. Mechanical Engineering & Automation, 2016, 5(2): 179–180.
28. Zhao T,Yu P, Miao Q, et al. One-dimensional single-photon position-sensitive silicon photomultiplier and its application in Raman spectroscopy [J]. Optics Express, 2017, 25(19): 22820–22828.
29. Takahashi M, Inome Y, Yoshii S, et al. A Technique for Estimating the Absolute Gain of a Photomultiplier Tube [J]. Nuclear Instruments & Methods in Physics Research, 2018, 894(10): 1–7.
30. Elorrieta M,Classen L, Reubelt J, et al. Characterisation of the Hamamatsu R12199-01 HA MOD photomultiplier tube for low temperature applications [J]. Journal of Instrumentation, 2019, 14(3):3015–3021.
31. Sun X, Kong M, Alkhazragi O, et al. Non-line-of-sight methodology for high-speed optical-wireless communication in highly turbid water [J]. Optics Communications, 2020, 461(1): 125264–125271.

32. Ning J, Gao G, Liang Z J, et al. Adaptive Receiver Control for Reliable High-speed Underwater optical-wireless communication with Photomultiplier Tube Receiver [J]. IEEE Photonics Journal, 2021, 13(4):1–7.
33. Wang H D. Analysis of current status and development trend of vacuum ultraviolet photomultiplier tubes [J]. Vacuum Electronics Technology, 2022, 4(2):23–28.
34. Yang W Z. Current status and future of long-wavelength detectors for optical communications [J]. Semiconductor Photonics, 1980, 2(3):45–61.
35. Watkins S E, Zhang C Z, Walser R M, et al. Laser-induced failure in biased silicon avalanche photodiodes [J]. Nist Special Publication, 1990, 12(1):95–104.
36. Li G Z. Analysis and optimum design on GexSi1-x/Si strained-layer superlattices avalanche photodetector [J]. Acta Optica Sinica, 1996, 16(6):17–24.
37. Ng J S, Tan C H, David J P R, et al. Effect of Impact Ionization in the InGaAs Absorber on Excess Noise of Avalanche Photodiodes [J]. IEEE Journal of Quantum Electronics, 2005, 41(8):1092–1096.
38. Guo J P, Liao C G, Wei Z G. Effect of reverse bias voltage on the penetration characteristics of avalanche photodiodes in Geiger mode [J]. Optics and Photonics, 2009, 7(5):9–11.
39. Chitnis D, Collins S. A flexible compact readout circuit for SPAD arrays [J]. Proc Spie, 2010, 7780(1):101–108.
40. Zhang Y, Liu B Q, Zhou B, et al. Signal to Noise Ratio Analysis of "cat's Eye" Target Active Detection System Based on Avalanche Photodiode [J]. Laser & Infrared, 2011, 41(11): 1240–1243.
41. Guo S, Ding Q X, Yang Y. Research on noise suppression technology of avalanche photodetectors [J]. Electro-Optics and Control, 2012, 19(3):69–73.
42. Adamo G, Agro D, Stivala S, et al. Measurements of Silicon Photomultipliers Responsivity in Continuous Wave Regime [J]. IEEE Transactions on Electron Devices, 2013, 60(11): 3718–3725.
43. Youn J S, Lee M J, Park K Y, et al. SNR characteristics of 850-nm OEIC receiver with a silicon avalanche photodetector [J]. Optics Express, 2014, 22(1):900–907.
44. Chen H T, Verbist J, Verheyen P, et al. High sensitivity 10Gb/s Si photonic receiver based on a low-voltage waveguide-coupled Ge avalanche photodetector [J]. Optics Express, 2015, 23(2): 815–822.
45. Goykhman I, Sassi U, Desiatov B, et al. On-Chip Integrated, Silicon–Graphene Plasmonic Schottky Photodetector with High Responsivity and Avalanche Photogain [J]. Nano Letters, 2016, 16(5):3005–3013.
46. Yin D, Yang X, He T, et al. InGaAs/InAlAs avalanche photodetectors integrated on silicon-on-insulator waveguide circuits [J]. Journal of Optical Technology, 2017, 84(5):350–357.
47. Juki T,Brandl P, Zimmermann H. Determination of the excess noise of avalanche photodiodes integrated in 0.35-μm CMOS technologies [J]. Optical Engineering, 2018, 57(4): 44101–44112.
48. Farrell A C,Meng X, Ren D, et al. InGaAs-GaAs nanowire avalanche photodiodes toward single-photon detection in free-running mode [J]. Nano Letters, 2019, 19(1):582–590.
49. Zhou X, Tan X, Lv Y, et al. High-uniformity 1×64 linear arrays of silicon carbide avalanche photodiode [J]. Electronics Letters, 2020, 56(17):895–897.
50. Eid M, Urooj S, Alwadai N M, et al. AlGaInP optical source integrated with fiber links and silicon avalanche photo detectors in fiber optic systems [J]. Indonesian Journal of Electrical Engineering and Computer Science, 2021, 23(2):847–854.
51. Li Z B, Li Y X, Ma Xu, et al. Measurement and suppression method of excess noise in avalanche photodiodes [J]. Infrared Technology, 2022, 44(4):343–350.
52. Lin Y F, Sun Y R, LU Z L. Design of double layer anti-reflection film for silicon PIN photodetector [J]. Journal of Hangzhou University, 1986(04):435–443.
53. Zhang J M, Conn D R. State-space modeling of the PIN photodetector [J]. Journal of Lightwave Technology. 1992, 1(5):10–18.

54. Zhang Y G, Cheng Z. SMA coaxially packaged high-speed photodetectors [J]. Semiconductor Optoelectronics, 1995, 2(1):23–30.
55. Jutzi M, Berroth M, Woehl G, et al. SiGe PIN photodetector for infrared optical fiber links operating at 1.25 Gbit/s [J]. Applied Surface Science, 2004, 224(1):170–174.
56. Giannopolous A V, Kasten A M, Hardy N, et al. Position Sensing Using an Integrated VCSEL and PIN Photodetector Microsystem [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2010, 76(3):15–25.
57. Shao H M, Zhang S L, Xie S, et al. Research of AlGaN/GaN PIN solar-blind ultraviolet photodetector with back-illumination [J]. Journal of Optoelectronics Laser, 2011, 22(7): 984–986.
58. Guan M J, Zhao Dong'e. Noise study based on PIN-type photoelectric conversion circuit [J]. Electronic Testing, 2012, 2(5):35–38.
59. Wu F, Li H Z, Du Ch M, et al. Characterization of optical threshold of PIN detector in free optical communication [J]. Modern Electronic Technology, 2014, 37(5):12–15.
60. Chaudhari P, Singh A, Topkar A, et al. Fabrication and characterization of silicon based thermal neutron detector with hot wire chemical vapor deposited boron carbide converter [J]. Nuclear Instruments & Methods in Physics Research, 2015, 779(11):33–38.
61. Rouse C, Zeller J W, Efstathiadis H, et al. Development of Low Dark Current SiGe Near-Infrared PIN Photodetectors on 300 mm Silicon Wafers [J]. Optics and Photonics Journal, 2016, 6(5):61–68.
62. Xu Z P, Jin C G, Yu Q, et al. Lidar reference light detection with PIN detectors [J]. Infrared and Laser Engineering, 2018, 47(10):267–273.
63. Shu B, Yan K, Zhong S. RLS denoising for photocurrent detection circuit of PIN-TIA photodetector [J]. Modern Electronics Technique, 2019, 24(1):1–4.
64. Wang N, Zhao B Q, Wang S, et al. Optoelectronic monolithic integration of PIN detector and transimpedance amplifier [J]. Infrared and Laser Engineering, 2021, 50(9):307–312.
65. Xiao F, Han Q, Ye H, et al. InP-based high-speed monolithic PIN photodetector integrated with an MQW semiconductor optical amplifier [J]. Japanese journal of applied physics, 2022, 61(1):12005–12013.
66. Shun Y. Precision active tracking technology using four-quadrant detector [J]. Laser & Infrared, 1979, (S1): 25–38.
67. Feng L G, Deng R L. Discussion on some problems in four-quadrant optoelectronic tracking technology [J]. Infrared and Laser Engineering, 1996, 25(1): 12005–12013.
68. Wang X, Xiu Q S, Liu G, et al. A Method for Improving the Detection Accuracy of the Spot Position of the Four-Quadrant Detector in a Free Space Optical Communication System [J]. Sensors, 2020, 20(24):7164–7173.
69. Kuang Z F, Feng Q B, Feng J Y, et al. Characterization of a four-quadrant detector for laser collimation [J]. Optical Technology, 2004, 12(4):387–389.
70. Wu Q B, Chen H Q, Liu Y. Research on noise detection technology of four-quadrant infrared photodetector [J]. Optics and Photonics, 2005, 28(5):28–31.
71. Zhao X, Tong S F, Jiang H L. Characterization of four-quadrant detectors [J]. Optical Precision Engineering, 2010, 18(10):2164–2170.
72. Yang G B, Zhang Z F, Zhai Y S, et al. Study on the effect of dead zone on the detection range and sensitivity of four-quadrant detectors [J]. Advances in Lasers and Optoelectronics, 2013, 50(6):156–161.
73. Schalkwyk S V, Meyer W E, Nel J M, et al. Implementation of an AlGaN-based solar-blind UV four-quadrant detector [J]. Physica B Condensed Matter, 2014, 439(1):93–96.
74. Tian Y M, Wang Y Z, Jiang Y J, et al. The Design of A focal Optical System of Beacon Tracking System in Inter-satellite Laser Communication [J]. Guidance & Fuze, 2016, 11(4): 54–57.
75. Grajower M, Desiatov B, Mazurski N, et al. Integrated on-chip silicon plasmonic four quadrant detector for near infrared light [J]. Applied Physics Letters, 2018, 113(14):143103–143110.

76. Wang X, Su X, Liu G, et al. Investigation of high-precision algorithm for the spot position detection for four-quadrant detector [J]. Optik, 2019, 203(6):163941–163950.
77. Liu P F, Song C L, Ma Y C. A four-quadrant detector design for dim light detection [J]. Foreign Electronic Measurement Technology, 2020, 39(5):99–104.
78. Safi H, Dargahi A, Cheng J. Beam Tracking for UAV-Assisted FSO Links With a Four-Quadrant Detector [J]. IEEE Communications Letters. 2021, 25(12): 3908–3912.
79. Sheng H M, Wang J K, Gao Q, et al. Noise characterization of GaAs/AlGaAs quantum well infrared detectors [J]. Semiconductor Technology, 1999, 5(1):26–29.
80. Zhao Y L, Li X J, Cai D M, et al. Dark current characterization of AlGaAs/GaAs multi-quantum well infrared detectors [J]. Infrared and Laser Engineering, 2006, 15(5):68–71.
81. Hill C J, Soibel A, Keo S A, et al. Mid-infrared quantum dot barrier photodetectors with extended cutoff wavelengths [J]. Electronics Letters, 2010, 46(18):1286–1271.
82. Chien L H. Quantum Dot Infrared Photodetectors Based on Structures with Potential Barriers: Modeling and Optimization [J]. Physica E Low Dimensional Systems & Nanostructures, 2011, 14(5):11–16.
83. Gustafsson O, Karim A, Berggren J, et al. Photoluminescence and photoresponse from InSb/InAs-based quantum dot structures [J]. Optics Express, 2012, 20(19):21264–21271.
84. Liu H M, Yang C H, Liu X, et al. Noise characterization of quantum dot infrared detectors [J]. Physics Letters, 2013, 62(21):454–459.
85. Kim J O, Ku Z, Krishna S, et al. Simulation and analysis of grating-integrated quantum dot infrared detectors for spectral response control and performance enhancement [J]. Journal of Applied Physics, 2014, 115(16): 101–112.
86. Srinivasan T, Mishra P, Jangir S K, et al. Molecular Beam Epitaxy growth and characterization of silicon—Doped InAs dot in a well quantum dot infrared photo detector (DWELL-QDIP) [J]. Infrared Physics & Technology, 2015, 70(1):6–11.
87. Karow M M, Munnelly P, Heindel T, et al. On-chip light detection using monolithically integrated quantum dot micropillars [J]. Applied Physics Letters, 2016, 108(8): 110–121.
88. Kim, Y.-H. Effects of imposed conditions and two types of electron transport on the noise of quantum dot infrared detectors [J]. Journal of Photonics, 2017, 46(4):81–86.
89. Wang Z. Research on dark current modeling of quantum dot-based infrared detectors [J]. Electronic Design Engineering, 2018, 26(23):89–93.
90. Baira M, Aljaghwani M, Salem B, et al. Investigation of GeSn/Ge quantum dots' optical transitions for integrated optics on Si substrate [J]. Results in Physics, 2019, 12(1):1732–1736.
91. Gansen E J, Nickel T B, Venner J M, et al. Sources of 1/f noise in QDOGFET single-photon detectors [J]. Physica E Low-dimensional Systems and Nanostructures, 2020, 118(1):61–70.
92. Gao X M, Xing T, Gao W Q, et al. Optical Communication Technology, 2022, (04):37–45.
93. Dai Y H, Ai Yong, Xiao Wei, et al. Research on balanced detectors for high-speed coherent optical communication [J]. Journal of Photonics, 2015, 44(1):179–185.
94. Guo L R, Hu Y H, Li Z, et al. Characterization of micro-Doppler effect laser balanced outlier detection [J]. Laser and Infrared, 2015, 45(3):247–251.
95. Shi Q Y, Ai Y, Liang H S, et al. Research and testing of balanced detectors in coherent optical communication [J]. Science, Technology and Engineering, 2016, 16(16):207–211.
96. Munshid M A. Design and Implementation Of Heterodyne Detection Based On Photonic Crystal Fiber Sensor [J]. Advances in Natural and Applied Sciences, 2017, 11(8):639–644.
97. Yong L, Yang A, Peng G, et al. Chromatic dispersion estimation based on heterodyne detection for coherent optical communication systems [J]. Optical Engineering, 2018, 57(1):11–16.
98. Zhang H F, Su B, Jones David R, et al. Design of a balanced detector in a high-speed terahertz time-domain spectroscopy system [J]. Journal of Terahertz Science and Electronic Information, 2019, 17(5):755–759.
99. Hu Q, Borkowski R, Schuh K, et al. Optical Field Reconstruction of Real-Valued Modulation Using a Single-Ended Photoreceiver with Half-Symbol-Rate Bandwidth [J]. Journal of Lightwave Technology, 2021, 39(4):1194–1203.

100. Huai Y C. Simulation Research of Double balanced Heterodyne Laser Detection System [J]. Software, 2013, 34(04):132–134.
101. Gao L, Wang C H, Li Y C et al. Balanced detection and reception experiment of 1.55um coherent laser wind radar [J]. Journal of Photonics, 2010, 39(6): 1064–1069.
102. Gao L, Wang C H, Li Y C, et al. Theoretical analysis of polarization mixing error in double-balanced outlier detection [J]. Infrared and Laser Engineering, 2010, 39(3):422–426.
103. Zhu J. Double Balanced Detection of Optical Fiber Acoustic Sensor Based on Sagnac Interferometer [J]. Electronic Science and Technology, 2013, 2(1):22–31.
104. Huai Y C. Simulation and study for dual-balanced heterodyne laser detection system based on the Opti System [J]. Software, 2013, 34(4):132–134.
105. Zheng F, Zhu G, Liu X, et al. Double balanced differential configuration for high speed InGaAs/InP single photon detector at telecommunication wavelengths [J]. Optoelectronics Letters, 2015, 11(2):21–30.
106. Peng C, Zhao C M, Zhang H Y, et, al. Simulation study on double-balanced coherent detection technology of laser wind radar [J]. Aerospace Return and Remote Sensing, 2015, 36(6):55–63.
107. Li Y, Zhang S, Zhou J, et, al. Research on the application of double-balanced detection in laser velocimetry [J]. Intense Laser and Particle Beam, 2016, 28(10):22–26.
108. Huang J, Zhu W. A new method for detecting ground fault detection in DC system based on double balanced bridge [J]. Electrical Measurement and Instrumentation, 2017, 54(8):75–79.
109. Liu H X, Zhou B, Gao Y C. Research on noise characteristics and influencing factors of APD detection system [J]. Laser Technology, 2018, 42(06):862–867.
110. Guo C L, Zeng D, Li J, et al. Noise model of photomultiplier tube [J]. Nuclear Electronics and Detection Technology, 2004(02):117–120.
111. Zhou H W. Design and Application research of four-quadrant detector noise test system [D]. Xi'an: Xidian University, 2013.
112. Chen M. Research on Channel Noise in Wireless Optical Coherent Communication and Its Influence on Coherent Detection Performance [D]. Xi'an: Xi'an University of Technology, 2018.

Noise Models in Optical Wireless Communication Systems

7

Contents

Introduction 206
Noise Distribution in Optical Wireless Communication Systems 206
Source Noise 208
  Research Progress 208
  Noise Models 209
Channel Noise 211
  Research Progress 211
  Noise Model 214
Beacon Noise 223
  Research Progress 223
  Noise Models 224
Performance Analysis of Optical Wireless Communication Systems Under the Influence of Turbulence 228
Conclusion 230
References 232

Abstract

Optical wireless communication has received widespread attention as a high-speed, high-bandwidth communication technology. However, noise is an important limiting factor in system performance and has a critical impact on optical wireless communication systems. This chapter reviews the research progress of noise modeling in optical wireless communication systems, summarizes the research results of Ke's model, the turbulence-noise model, thermal-noise model, background optical-noise model, etc., and further analyzes the performance of systems under the influence of turbulence. Finally, the shortcomings of current research are highlighted, and future development directions are discussed.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_7

Keywords

Optical wireless communication · Ke's model · Turbulence noise model · Noise modeling

Introduction

With the increasing demand for wireless communication and the growing scarcity of spectrum resources, optical wireless communication systems have attracted widespread attention as an emerging high-speed, high-bandwidth communication technology. Optical wireless communication systems use light as a carrier for data transmissions. This has the advantages of high-capacity, low-latency, and interference resistance [1] and is widely used in high-density communication scenarios and special environments.

However, noise is a nonnegligible factor in optical wireless communication systems and has a critical impact on system performance. Noise sources include turbulence noise, background optical noise, and thermal noise. These noises introduce additional interference into the transmission, reception, and processing of optical signals and reduce the signals' quality and reliability. Therefore, the study of noise and noise modeling in optical wireless communication systems is critically important for improving the performance of communication systems.

This chapter reviews the research progress in noise modeling in optical wireless communication systems, outlines the current research results in noise modeling, and analyzes the system performance under the influence of noise. Finally, it summarizes the shortcomings of current research and discusses future development directions for noise-modeling research. This review improves the understanding of how noise affects system performance and provides guidance and insights for noise suppression and system optimization.

Noise Distribution in Optical Wireless Communication Systems

Optical wireless communication is a broadband wireless transmission and access technology that combines the optical and wireless communication technologies. It uses light waves as carriers and the atmosphere, water, or outer space as channels for information transfer. It is also known as free-space optical (FSO) communication.

In optical wireless communication, the information output by the source is first modulated to an optical carrier, and the carrier (light beam or field) is launched through an optical channel (atmosphere) using a transmitter. The optical transmitter includes source coding, channel coding, modulation, optical-signal amplification, and a transmitting antenna. The optical signal is then optically collected and processed at the receiver end. The optical receiver consists of an optical-signal collection antenna, spatial optical-fiber coupling unit, preamplifier, detector, and demodulator [2].

Typically, various noises are present in the optical-field transmission process, which causes the performance of the communication system to deteriorate. Therefore, it is crucial to model and analyze these noises. In this study, we individually analyze all types of typical noise introduced in the source, channel, and beacon of an optical wireless communication system. A typical optical wireless communication system and its corresponding noise are shown in Fig. 7.1.

Optical wireless communication typically uses lasers as light sources. The laser itself introduces noise, mainly laser-intensity and phase noise. Laser-intensity noise is caused by the randomness of the photons generated within the laser, which can lead to fluctuations in the optical power. Laser phase noise is mainly caused by factors such as the natural linewidth, technical noise, and environmental interference, which may affect the stability of the optical signal.

The channels of optical wireless communication systems include atmospheric, indoor, ultraviolet-light scattering, and underwater channels. Channel noise mainly includes turbulence, phase, aiming, and background-light noise. Turbulence noise is caused by turbulence phenomena in the channel, which can cause the received signal to fade or distort. Phase noise is mainly produced by changes in the refractive index of light, caused by factors such as temperature and humidity in the channel, resulting in phase distortion and timing jitter of the received signal.

Aiming noise is caused by the inaccurate orientation of an optical receiver, which can cause changes and fluctuations in the intensity of the received signal, affecting its

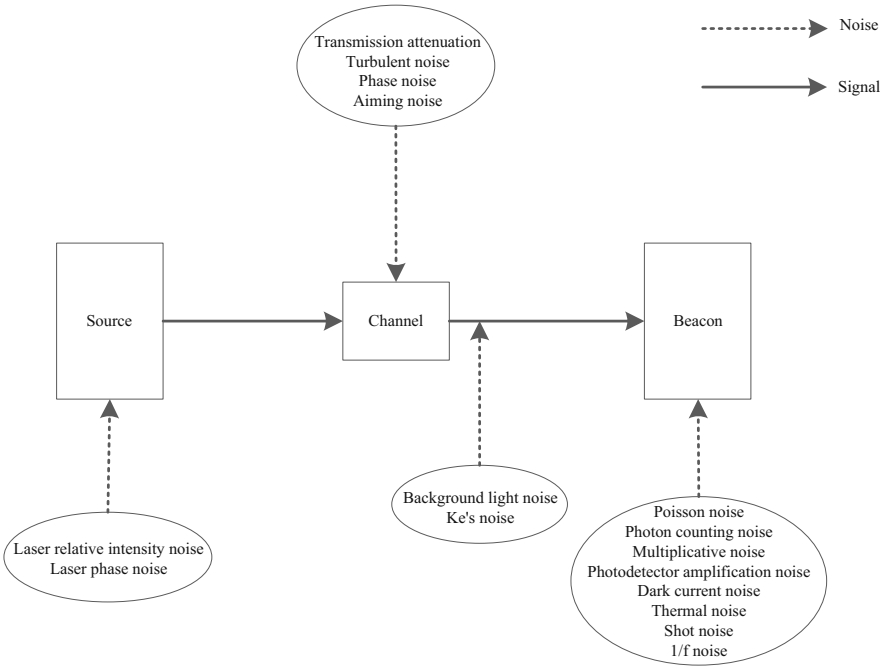


Fig. 7.1 Noise distribution in an optical wireless communication system

stability and reliability. Background-light noise mainly originates from the noise caused by other light sources in the environment surrounding the optical communication system. It can reduce the signal-to-noise ratio, leading to a decrease in the quality of the received signal.

A sink in an optical wireless communication system refers to a PIN or avalanche photodiode. Beacon noise includes Poisson, photon-counting, multiplication, thermal, dark-current, photodetector-amplification, shot, and 1/f noise. These noises primarily originate from the photodetector, which limits the sensitivity and reliability of the system.

Studying noise models is important precisely because noise has such a negative impact on optical wireless communication systems. By analyzing noise models, we can gain a deeper understanding of the impact of noise on communication performance. This helps evaluate the reliability and stability of a system and provides a foundation for optimizing the system performance.

Source Noise

Research Progress

It is an undeniable factor that source noise has a significant impact on the performance of the entire system. Source noise originates from the nonideal properties of optical emission devices, such as the spontaneous emission of lasers and other nonlinear effects, which can cause noise in the signal source and interfere with the transmission quality.

In recent years, researchers have made many achievements in the study of source-noise models. In 2003, Lu [3] analyzed and compared the characteristics of solid-state and semiconductor lasers in free-space optical communication systems. In 2004, Yang [4] found that a CO_2 laser, as a light source for optical wireless communication, could better adapt to harsh atmospheric environments and improve the system performance. In 2005, Gensty et al. [5] proposed a semiclassic noise model for the relative-intensity noise of quantum-cascade lasers.

In 2006, Yu et al. [6] designed a new type of low-noise fiber laser using a composite-cavity structure, which reduced the noise by approximately 25 dB at the 90-kHz relaxation oscillation peak. In 2008, Camatel et al. [7] evaluated two novel techniques for measuring laser phase noise in coherent transmission systems using a homodyne optical phase-locked loop and a signal-source analyzer. In 2010, Di Domenico et al. [8] proposed an approximate formula for evaluating laser linewidths that could be applied to any noise spectral density.

In 2011, Xie et al. [9] simulated and tested the intrinsic correlation between laser phase noise, laser coherence length, and reflected signal intensity. The results indicated that when the testing distance was close to the coherent length with a strong reflection signal in the link, the impact of the laser phase noise was more severe, and the impact range also increased. In 2012, Liu et al. [10] studied the impact of laser phase noise on the modulation-format selection, parameter

optimization, and carrier-phase estimation performance of coherent fiber-optic communication systems. The results indicated that when the carrier-phase estimation method was used in coherent fiber-optic communication systems, the optimal length of the average symbol sequence in the algorithm was inversely proportional to the laser phase noise.

In 2016, Chen et al. [11] derived an expression for the nonlinear intermodulation distortion of lasers using Bessel functions and investigated its effect on the system-carrier intermodulation noise ratio and bit-error rate (BER). In 2017, Warid et al. [12] theoretically investigated the static characteristics, relative-intensity noise, and intensity-modulation response of quantum-cascade lasers and analyzed the noise of such semiconductor-laser spectra and the modulation response of such semiconductor lasers. The results showed that static characteristics, such as the threshold injection current and steady-state photon density, deviated from the ideal values as the optical-phonon emissivity increased.

In the same year, Song et al. [13] used power and integral spectra to qualitatively analyze the phase noise of fiber lasers with different pump powers, seed-source powers, and wavelengths. The results showed that the higher the pump power, the greater the proportion of low-frequency phase noise; for different seed-source powers, the proportion of phase noise in the tested frequency bands was basically at the same level. For the wavelengths of the different seeds tested, the phase noise in the high proportion of the frequency band was different.

In 2019, Bai [14] derived the relationship between the laser linewidth, frequency-noise power-spectral density, and phase-noise power-spectral density and derived a method to calculate the laser linewidth by measuring the laser phase noise. In 2021, Didier et al. [15] investigated the relative-intensity noise in 4.1-μm multimode interband cascade lasers. The results indicated that the relaxation frequency increased with the bias current at much higher than the threshold current. The simulation results were in agreement with the experimental results.

In 2022, Du [16] et al. numerically investigated the conditions for the codriven chaotic synchronization of distributed Bragg-reflector (DBR) lasers, effect of parameter mismatches, and chaotic-synchronization switch keying. In 2022, Seminara et al. [17] investigated a free-space quantum-cascade laser-based method and carried out a detailed experimental analysis of the noise conditions that could occur in an optical communication system. The results showed that there were two noise regions in the system corresponding to different channel-attenuation conditions, and error-free communication could be achieved within an effective distance of 8 km with 1 km low-visibility conditions.

Noise Models

Laser Relative-Intensity Noise Model

In a communication system, the intrinsic noise of the laser is the main source of source noise, which can be expressed in terms of the relative-intensity noise (RIN). When the average power of the emitted beam is P, and it is received by a detector

with responsivity ρ and bandwidth Δf, the average value of the current in the detector is equal to the product of the responsivity and the average power, ρP [9]; then, the mean-square value of the current is

$$\overline{i_{NL}^2} = \mathrm{RIN}(\rho P)^2 \Delta f. \tag{7.1}$$

The average power which caused fluctuations in the light intensity of the laser is

$$\overline{P_{NL}^2} = \frac{\overline{i_{NL}^2}}{\rho^2}. \tag{7.2}$$

The relative-intensity noise of the laser can be obtained as

$$\mathrm{RIN} = \frac{\overline{P_{NL}^2}}{P^2 \Delta f}, \tag{7.3}$$

where $\overline{i_{NL}^2}$ is the mean-square value of the current and $\overline{P_{NL}^2}$ is the average power. Thus, it can be seen that the laser noise is mainly related to the average beam power and detector bandwidth.

Laser Phase-Noise Model

Spontaneous-radiation noise can cause phase fluctuations in the laser, resulting in a nonzero spectral linewidth Δv. In semiconductor lasers, the linewidth Δv is generally large, up to megahertz, because a small number of photons is confined in a small cavity, and the linewidth-enhancement factor is not negligible.

The relationship between the spectrum of the radiated light field $E(t) = \sqrt{P}\exp(j\phi)$ and the field autocorrelation function $\Gamma_{EE}(t)$ is

$$S(\omega) = \int_{-\infty}^{\infty} \Gamma_{EE}(t) e^{-j(\omega-\omega_0)\tau} \mathrm{d}\tau \tag{7.4}$$

$$\Gamma_{EE}(t) = \langle E^*(t)E(t+\tau)\rangle = \langle \exp(j\Delta\phi(t))\rangle = \exp\left(-\left\langle \Delta\phi^2(t)\right\rangle/2\right), \tag{7.5}$$

where $\Delta\phi(t) = \phi(t+\tau) - \phi(t)$, E represents the light field, t represents time, τ represents the time interval, j represents an imaginary number, ω represents the angular frequency, and ϕ represents the phase.

The Wiener process can be used to describe the phase noise process of lasers [10], namely,

$$S_{EE}(v) = \frac{P}{2\pi\Delta v}\left[\frac{1}{1+\left(\frac{2(v+v_0)}{\Delta v}\right)^2} + \frac{1}{1+\left(\frac{2(v-v_0)}{\Delta v}\right)^2}\right], \tag{7.6}$$

where P is the optical power, v is the frequency, v_0 is the center frequency, and v is the laser linewidth.

Channel Noise

Research Progress

When a light beam propagates in a channel, the transmitted beam may attenuate or drift, owing to the influence of particulate molecules and meteorological factors in the channel. To ensure the quality and reliability of the wireless communication system, the noise introduced in it must be analyzed.

In 2005, Razavi et al. [18] investigated the effects of atmospheric turbulence, background light, amplified spontaneous radiation, and receiver thermal noise. In the same year, Kiasaleh et al. [19] analyzed a weak turbulence (clear-air) scenario and a negative-exponential distribution of the received signal strength and provided numerical results for the pulse-position modulation case to elucidate the effect of turbulence on the overall performance.

In 2008, Zhu [20] conducted an in-depth investigation on the performance of multiple modulation modes for indoor optical wireless communication and investigated the high-speed transmission process in which a root-mean-square (RMS) delay was generated by multipath channels and the resulting intercode interference. In the same year, Mabaoke et al. [21] explored the characteristics of laser-propagation attenuation in fog and analyzed the effects of field-of-view angle variations on the signal-to-noise ratio.

Jaruwatanadilok et al. [22] used the vector radiative-transfer theory to model an underwater optical wireless communication channel and quantified the scattering effect as a function of the distance and bit-error rate using the Monte Carlo method. In 2009, Cochenour et al. [23] investigated the effects of underwater scattering on optical communication signals, received optical power, and the link range. In 2010, Meng [24] used the classic quadratic generalized Lagrange method to analyze the optical wireless communication model under equation constraints. They established an optical wireless communication model affected by the atmospheric-transmission window, as well as atmospheric turbulence.

In 2011, Wang [25] constructed a mathematical computational model to evaluate the error performance of optical wireless communication systems under weak turbulence. In 2012, Tu [26] proposed a method for calculating the bit-error rate under binary amplitude-keying modulation based on a hybrid noise model. The simulation results showed that the performance of free-space optical communication was significantly affected by atmospheric turbulence and noise. In the same year, Mohamed et al. [27] presented the impact of inclement weather conditions, such as rain, fog, and snow, on the transmission performance of optical wireless communication systems.

In 2013, Yao et al. [28] analyzed a channel model of underwater optical wireless communication (UOWC) and derived the BER expression of fixed-length digital

pulse-spacing modulation in an underwater optical wireless communication model. In the same year, Tang et al. [29] proposed a model of the channel impulse response with a closed expression for the double-gamma function. Ansari [30] conducted an in-depth study of the aiming-error problem in optical communication systems using the Meijer G-function and integrated the results into a unified expression.

In 2014, Nistazakis et al. [31] studied the wireless BER performance of optical communication systems over atmospheric-turbulence channels modeled with a gamma–gamma distribution or negative-exponential distribution. In the same year, Varotsos [32] analyzed the effects of group velocity dispersion and atmospheric turbulence on communication systems, owing to laser atmospheric propagation. Zhu et al. [33] proposed a phase-noise measurement scheme based on the delay-line method. It employed a multifunctional microwave photonic processor that simultaneously performed an electro-optical conversion, provided a time delay, and controlled the phase of the output microwave signal.

In 2015, Ke [34] established a space optical wireless communication receiving-diversity system model under strong turbulence. He verified its applicability and antijamming capability in an atmospheric environment by simulation. In the same year, Yang [35] derived an expression for the false-symbol rate of an optical wireless communication system under multipulse-position modulation using the exponential Weber fading model. Rashed [36] developed a model for evaluating the operational performance of an optical wireless communication system under different fog-density levels and in the presence of noise. Xu et al. [37] proposed a model for evaluating the effect of balanced augmented phase noise on the performance of coherent optical transmission systems.

In 2016, Wu et al. [38] analyzed atmospheric-turbulence channel-intensity fluctuation based on the K-distribution and numerically analyzed and compared the BER of circularly polarized keying with that of switched keying. The results showed that the system BER increased with an increase in the atmospheric scintillation index at the same signal-to-noise ratio, whereas circularly polarized keying modulation outperformed switched-keying modulation in strongly turbulent channels.

In 2016, Hamza et al. [39] investigated the effect of ambient noise induced by underwater solar radiation on UOWC from the perspective of the BER. This study was conducted on different types of photodetectors to highlight the practical limitations of establishing UOWC links in the presence of undersea solar noise and provided valuable insights into the design of UOWC links. In 2017, Anbarasi et al. [40] used relaying, diversity, and modulation techniques in different channels to mitigate turbulence in FSO systems and compared their performances.

In 2017, Fan et al. [41] modeled sky background-optical noise and performed theoretical-analysis and simulation studies. The simulation results showed that the sky background light decreased the signal-to-noise ratio of the detectors used in the coarse and fine tracking units, which led to a decrease in tracking accuracy and an increase in the communication BER. In 2017, Zhang [42] described the specific classification of noise in free-space optical communication systems and time-varying fading and stationary fading models owing to turbulence.

In 2018, Ke [43] considered an optical wireless subcarrier multivariate phase-shift keying modulation system and investigated the effect of the demodulation-module phase noise on the error performance of the optical wireless communication system. He also simulated and analyzed the effects of the light-intensity scintillation variance, phase noise, and modulation order on the error performance of the system.

In 2019, Wang [44] derived a Fourier-series expression for the probability-density function of the phase-noise model of a Tikhonov distribution. In 2020, Xing [45] analyzed the effects of absorption, scattering, and solar-radiation noise on the performance and power distribution of underwater optical communication systems.

In the same year, Hong et al. [46] used Monte Carlo simulation to construct an underwater optical wireless communication model for long-distance weak links and investigated the effects of the demodulation threshold, signal optical power, and noise optical power on the system performance. The experimental results showed that the simulation results of the established system model were basically consistent with the experimental data.

Fu et al. [47] conducted a simulation study of a UOWC system to investigate the effects of the beamwidth, receiving aperture, ocean-turbulence parameters, and aiming error on the performance of the average BER and interruption probability. Wang et al. [48] conducted a simulation study of the low signal-to-noise ratio of the receiving spot through digital signal filtering and denoising to improve the spot-position detection accuracy. The spot-position detection error of this method in a high background-noise environment was only 0.0277 mm, and the root-mean-square error was 0.0065 mm.

Liu [49] analyzed the effect of the absorption scattering of seawater channels on the attenuation of light and compared the time-domain expansion of a light pulse under different water-quality conditions by simulation. Wang [50] modeled the underwater optical propagation characteristics and analyzed the attenuation due to light absorption and scattering, background noise, and energy consumption in the ocean when transmitting data. Zhang et al. [51] proposed an asymptotic-growth metalearning adaptive-multirate scheme based on orthogonal frequency-division multiplexing and analyzed it over UOWC turbulent channels.

In 2021, Ke et al. [52] investigated the effect of atmospheric turbulence on optical wireless communications, exploring phenomena such as beam spreading, drifting, and scintillation. In the same year, Abigail et al. [53] designed a receiver system based on a deep-confidence network. The simulation results showed that the system performance improved by 13.2 dB at a 10^{-3} BER.

Xu [54] used lognormal and gamma–gamma channel models to describe weak to strong turbulence. In addition, the average BER and channel capacity of the FSO communication system under the influence of background, thermal, and quantum noises were analyzed. Han [55] analyzed the channel characteristics and noise model of an indoor visible-light communication system, focusing on the effects of scattered-particle and thermal noise on system performance.

Ke et al. [56] proposed a background optical-noise model for nighttime visible-light communication systems in vehicular networks, based on the superposition of double-Gaussian functions. Majlesein [57] introduced an analytical model of

scattering noise and investigated its effect on the BER performance of UOWC systems for three types of water and a range of transmission-link spans.

In 2022, Zhang et al. [58] analytically modeled the receiver-side noise of a UOWC system and built an experimental UOWC platform with an 852 nm wavelength under ambient-light noise interference. They demonstrated that the use of optical filters could effectively improve the communication performance and link length of underwater optical wireless communication under ambient-light noise interference.

In the same year, Li et al. [59] systematically studied and compared the theory and construction of various atmospheric-turbulence spectral models and detailed the atmospheric-turbulence models used in practical studies. Chen et al. [60] proposed an architecture for wavelength-division-multiplexed coherent data-center links and found that in each frequency comb, the phase noise in all the combs was completely dominated by the seed laser's optical phase noise and the phase noise of the microwave oscillator.

Yang et al. [61] developed a composite channel model under the synergistic effect of ocean turbulence and fading channels and analyzed the system performance by simulating the signal waveforms and counting the BERs. In 2023, Ke et al. [62] derived a multiplicative-noise model for atmospheric turbulence and conducted experimental measurements. The results showed that the effect of turbulent-multiplicative noise on the communication quality of optical communication intensified in the order of sunny, hazy, snowy, light, medium, and heavy rain, and cloudy days. The skewness and steepness of the noise signals increased gradually, and the fading probability of the light intensity of the signals increased.

In the same year, Jin et al. [63] investigated and discussed the performance of broadband long-haul nonlinear fiber-optic communication systems under balanced enhanced phase noise. The results showed that the growth of the symbol rate and transmission distance exacerbated the distortion of the C-band system. In 2023, Li et al. [64] discussed optical design, driving circuitry, and channel modeling in visible-light full-duplex communication.

Noise Model

Transmission-Attenuation Model

When a laser propagates in a channel, owing to the existence of a variety of gas molecules and particles in the channel, such as dust, smoke, wind, rain, snow, and other meteorological changes, a portion of the optical radiation energy is absorbed and transformed into other forms of energy (such as heat). The other part of the energy is scattered and deviates from the original propagation direction.

The total effect of the absorption and scattering is the attenuation of the transmitted optical-radiation intensity. The expression is

$$I_l = \exp[-c(\lambda)L], \tag{7.7}$$

where $c(\lambda)$ is the attenuation coefficient and L is the optical-transmission distance of the signal.

Turbulent-Noise Model

Turbulent noise in the channel originates from the scintillation effect. In addition, some particles, such as aerosols, sand, and dust, move randomly in the presence of the wind field, which leads to a random scintillation of the light intensity of the received signal. Based on field measurements, the received light intensity exhibits a specific probability-density distribution pattern in both cases. Therefore, this distribution can be used as a statistical feature to analyze the optical-signal transmission process.

In the field of optical wireless communication, the light-intensity distribution of the channel in weak-, medium-strong, and strong-turbulence scenarios can be described by probabilistic models, such as the lognormal, gamma–gamma, and K-distribution models [33].

Lognormal Model

The lognormal model is suitable for describing light-intensity flickering under weak turbulence. Under weak-turbulence conditions, fluctuations in the optical-signal intensity are mainly affected by random fluctuations in the refractive index of the atmosphere. The lognormal model assumes that the natural logarithm of the optical-signal intensity follows a normal distribution. The model is simpler but performs poorly when describing medium-strong and strong turbulence.

The probability-density function of the lognormal model can be expressed as

$$f_I(I) = \frac{1}{2I}\frac{1}{\sqrt{2\pi\sigma_\chi^2}}\exp\left\{-\frac{\left[\ln\left(\frac{I}{I_0}\right)+2\sigma_\chi^2\right]^2}{8\sigma_\chi^2}\right\}, \tag{7.8}$$

where I represents the received light intensity, I_0 represents the received light intensity before turbulence, and σ_χ^2 represents its logarithmic variance.

Γ-Γ Model

The Γ-Γ model, which is obtained by multiplying two independent gamma distributions, provides a better description of the fluctuations of the optical-signal intensity under moderate to strong turbulence. It is affected by multiple factors, such as random fluctuations of the atmospheric refractive index and scattering.

The probability-density function of the Γ-Γ model can be expressed as [39]

$$f_I(I) = \frac{2(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)} I^{\frac{\alpha+\beta}{2}-1}\mathrm{K}_{\alpha-\beta}\left(2\sqrt{\alpha\beta I}\right), \tag{7.9}$$

where $\Gamma(\cdot)$ denotes the gamma function and $K_n(\cdot)$ is the n-order Type-II modified Bessel function. α and β represent the large-scale and small-scale eddy-current characteristics, respectively, and are usually calculated by the following equations:

$$\begin{aligned}\alpha &= \left[\exp\left(\frac{0.49\sigma_l^2}{\left(1+0.65d^2+1.11\sigma_l^{12/5}\right)^{7/6}}\right)-1\right]^{-1}\\ \beta &= \left[\exp\left(\frac{0.51\sigma_l^2\left(1+0.69\sigma_l^{12/5}\right)^{-5/6}}{\left(1+0.9d^2+0.62d^2\sigma_l^{12/5}\right)^{5/6}}\right)-1\right]^{-1},\end{aligned} \tag{7.10}$$

where $\sigma_l^2 = 1.23C_n^2k^{7/6}L^{11/6}$ is the Rytov variance, $d = (kD^2/4L)^{1/2}$, $k = 2\pi/\lambda$ is the number of light waves, λ is the wavelength, D is the receiver-aperture diameter, and L is the beam-transmission distance. C_n^2 is the atmospheric refractive-index structure constant. The light-intensity scintillation index σ_I^2 for the Γ-Γ model is

$$\sigma_I^2 = \exp\left[\frac{0.49\sigma_l^2}{\left(1+0.65d^2+1.11\sigma_l^{12/5}\right)^{7/6}}+\frac{0.51\sigma_l^2\left(1+0.69\sigma_l^{12/5}\right)^{-5/6}}{\left(1+0.9d^2+0.62d^2\sigma_l^{12/5}\right)^{5/6}}\right]. \tag{7.11}$$

K-Distribution Model

Under strong turbulence, the fluctuation of optical-signal intensity has complex characteristics consisting of small- and large-scale fluctuations, and the K-distribution model can well describe the optical-intensity scintillation in this case.

The probability-density function of the K-distribution model is expressed as follows:

$$f_I(I;\alpha_K) = \frac{2\alpha_K}{\Gamma(\alpha_K)}(\alpha_K I)^{(\alpha_K-1)/2}K_{\alpha_K-1}\left(2\sqrt{\alpha_K I}\right), \tag{7.12}$$

where α_K is a parameter related to the variance of the received light intensity ${\sigma_I}^2$; its relationship with ${\sigma_I}^2$ is

$$\alpha_K = \frac{2}{{\sigma_I}^2-1}. \tag{7.13}$$

Negative-Exponential Distribution Model

In the case of strong turbulence, the radiation field of the light wave can be approximated as a Gaussian distribution with zero mean; therefore, the

light-intensity distribution is approximated as an exponential distribution, which can be expressed as

$$f(I) = \frac{1}{I_0} \exp\left(-\frac{I}{I_0}\right), \tag{7.14}$$

where $I_0 = E(I)$.

Malaga Distribution Model

The Malaga turbulence model describes several different types of turbulence-intensity distribution models, including the lognormal, gamma–gamma, K, and negative-exponential distributions. The probability-density function of the Malaga distribution for the turbulence-channel light intensity is defined as

$$f_I(I) = A \sum_{k=1}^{\beta} a_k I^{\frac{\alpha+k}{2}-1} K_{\alpha-k}\left(2\sqrt{\frac{\alpha\beta I}{\gamma\beta + \Omega'}}\right), \tag{7.15}$$

where

$$A = \frac{2\alpha^{\frac{\alpha}{2}}}{\gamma^{1+\frac{\alpha}{2}}\Gamma(\alpha)} \left(\frac{\gamma\beta}{\gamma\beta + \Omega'}\right)^{\beta+\frac{\alpha}{2}} \tag{7.16}$$

$$a_k = \binom{\beta - 1}{k - 1} \frac{(\gamma\beta + \Omega')^{1-\frac{k}{2}}}{(k-1)!} \left(\frac{\Omega'}{\gamma}\right)^{k-1} \left(\frac{\alpha}{\beta}\right)^{\frac{k}{2}} \tag{7.17}$$

$$\Omega' = \Omega + 2\rho b_0 + 2\sqrt{2b_0\Omega\rho}\cos(\varphi_A - \varphi_B), \tag{7.18}$$

where α is a positive real parameter related to the effective number of the large-scale turbulence, and β is a natural number indicating the number of fading parameters. Ω' is the coherent mean optical power, γ is the receiver signal-to-noise ratio, ρ is the ratio of the scattered optical power coupled to the line of sight to the total optical power, and φ_A and φ_B are the phases of the line-of-sight propagation component and the propagation-coupling component of the light wave in the atmospheric-turbulence channel, respectively.

Atmospheric-turbulence channel models, such as the lognormal, gamma–gamma, and K-distribution models, can be obtained by setting the corresponding α and β parameters to the Malaga-distributed light-intensity probability-density function.

Exponential-Generalized Gamma Distribution Model

The exponential-generalized gamma distribution model is a combination of the exponential and generalized gamma distributions, which is more flexible and can characterize turbulence phenomena from weak to strong. It also considers the turbulence characteristics when air bubbles and temperature and salinity gradients are present in the channel (Table 7.1):

Table 7.1 Turbulence-channel model classification diagram

Channel model	Probability-density function	Adapted to turbulence range
Lognormal model	$f_I(I) = \frac{1}{2I}\frac{1}{\sqrt{2\pi\sigma_\chi^2}}\exp\left\{-\frac{\left[\ln\left(\frac{I}{I_0}\right)+2\sigma_\chi^2\right]^2}{8\sigma_\chi^2}\right\}$	Weak turbulence
Γ-Γ model	$f_I(I) = \frac{2(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)} I^{\frac{\alpha+\beta}{2}-1}\mathrm{K}_{\alpha-\beta}(2\sqrt{\alpha\beta I})$	Medium to strong turbulence
K-distribution model	$f_I(I;\alpha_\mathrm{K}) = \frac{2\alpha_\mathrm{K}}{\Gamma(\alpha_\mathrm{K})}(\alpha_\mathrm{K} I)^{(\alpha_\mathrm{K}-1)/2}\mathrm{K}_{\alpha_\mathrm{K}-1}\left(2\sqrt{\alpha_\mathrm{K} I}\right)$	Strong turbulence
Negative-exponential distribution model	$f(I) = \frac{1}{I_0}\exp\left(-\frac{I}{I_0}\right)$	Strong turbulence
Malaga distribution model	$f_I(I) = A\sum_{k=1}^{\beta} a_k I^{\frac{\alpha+k}{2}-1} K_{\alpha-k}\left(2\sqrt{\frac{\alpha\beta I}{\gamma\beta+\Omega'}}\right)$	Weak, medium, and strong turbulence
Exponential-generalized gamma-distribution model	$f_I(I) = \frac{\omega}{\lambda}\exp\left(-\frac{I}{\lambda}\right) + (1-\omega)c\frac{I^{ac-1}\exp\left(-\left(\frac{I}{b}\right)^c\right)}{b^{ac}\Gamma(a)}$	Weak, medium, and strong turbulence

$$f_I(I) = \omega f(I;\lambda) + (1-\omega)g(I;[a,b,c]) \tag{7.19}$$

$$f(I;\lambda) = \frac{1}{\lambda}\exp\left(-\frac{I}{\lambda}\right) \tag{7.20}$$

$$g(I;[a,b,c]) = c\frac{I^{ac-1}\exp\left(-\left(\frac{I}{b}\right)^c\right)}{b^{ac}\Gamma(a)}. \tag{7.21}$$

Therefore,

$$f_I(I) = \frac{\omega}{\lambda}\exp\left(-\frac{I}{\lambda}\right) + (1-\omega)c\frac{I^{ac-1}\exp\left(-\left(\frac{I}{b}\right)^c\right)}{b^{ac}\Gamma(a)}. \tag{7.22}$$

The model is weighted using the exponential and generalized gamma distributions. $f(I;\lambda)$ is an exponential-distribution function; $g(I;[a,b,c])$ is a generalized gamma-distribution function; ω is the mixed weight coefficient of the distribution, which satisfies $0 < \omega < 1$; λ is a parameter of an exponential distribution; a, b, and c are parameters of the generalized gamma distribution; and $\Gamma(\cdot)$ is a gamma function.

Phase-Noise Model

Phase variations caused by fluctuations in the refractive index of atmospheric turbulence lead to phase noise. In the communication process, the photodetector converts the optical signal at the receiving end into an electrical signal, which is subsequently multiplied by the signal output from the local oscillator to obtain an intermediate-frequency signal, which is coherently demodulated. However, with the passage of time, the phase of the output signal of the local oscillator may fluctuate,

and the noise caused by such fluctuations may cause difficulties in demodulating the subcarrier signal, which may lead to the deterioration of the system performance.

The phase noise φ generated by the local oscillator follows the Tikhonov distribution [40], and its probability-density function is

$$P(\varphi) = \frac{\exp\left[\frac{\cos(\varphi)}{\sigma_\varphi{}^2}\right]}{2\pi \mathrm{I}_0\left(\frac{1}{\sigma_\varphi{}^2}\right)}, |\varphi| \le \pi, \tag{7.23}$$

where $\sigma_\varphi{}^2$ is the phase-noise variance, and $\mathrm{I_n}()$ is the first type of deformed Bessel function of order n.

The Jacobi–Anger formula can convert Eq. (7.23) into

$$\exp\left[\frac{\cos(\varphi)}{\sigma_\varphi{}^2}\right] = \mathrm{I}_0\left(\frac{1}{\sigma_\varphi{}^2}\right) + 2\sum_{n=1}^{\infty} \mathrm{I_n}\left(\frac{1}{\sigma_\varphi{}^2}\right)\cos n\varphi. \tag{7.24}$$

Substituting Eq. (7.24) into Eq. (7.23), the Fourier series can be used to represent the probability-density function of the phase noise as

$$p(\varphi) = \frac{1}{2\pi} + \sum_{n=1}^{\infty} c_n \cos(n\varphi), \tag{7.25}$$

where $c_n = \frac{\mathrm{I}_n\left(\frac{1}{\sigma_\varphi{}^2}\right)}{\pi \mathrm{I}_0\left(\frac{1}{\sigma_\varphi{}^2}\right)}$ is the Fourier coefficient of the Tikhonov phase-noise probability-density function.

Aiming-Noise Model

When a light beam is transmitted through a channel, it causes a misalignment between the receiver and transmitter, owing to phenomena such as beam expansion and spot drift, which, in turn, leads to aiming errors in the system.

Assuming that a beam with beamwidth w_z arrives at a receiver that is distance z from the transmitter, the radius of the receiver aperture is R, and the attenuation coefficient due to the aiming error can be expressed as [44]

$$h_p(r, \mathrm{z}) = A_0 \exp\left(-\frac{2r^2}{w^2{}_{z_{eq}}}\right) \tag{7.26}$$

$$A_0 = [\mathrm{erf}(v)]^2. \tag{7.27}$$

$$v = \frac{\sqrt{\pi}\mathrm{R}}{\sqrt{2}\mathrm{w}_z} \tag{7.28}$$

$$w_{Z_{eq}}^2 = \frac{\mathrm{w}_z^2 \sqrt{\pi}\mathrm{erf}(v)}{2v\exp(-v^2)}, \tag{7.29}$$

where h_p is the loss owing to the aiming error, r is the radial deviation, z is the distance between the transmitter and receiver, A_0 is the $r = 0$ received power component, erf(·) is the error function, v is the beam-transfer rate, w_z is the beam width, and $w_{Z_{eq}}^2$ is the equivalent beamwidth. According to the Rayleigh distribution, the probability-density function of h_p is

$$f_{h_p}(h_p) = \frac{\gamma^2}{A_0^{\gamma^2}} h_p^{\gamma^2 - 1}, 0 \le h_p \le A_0, \tag{7.30}$$

where $\gamma = w_{z_{eq}}/2\sigma_s$ represents the ratio of the equivalent beam radius at the receiving end to the standard deviation of the aiming-error displacement and σ_s^2 represents the equivalent variance at the receiving end (Fig. 7.2).

In summary, it can be seen that in an optical wireless communication system, the impact of the aiming error on the system is mainly reflected as follows: By choosing the appropriate a/w_z ratio, the f_{h_p} curve is an incremental function of the relationship, which can improve the performance of the communication system; otherwise, it will increase the signal-fading probability.

Background-Light Noise Model

Interfering light emitted or reflected from light sources other than the signal light, such as sunlight, moonlight, and artificial light sources, and received by the optical wireless receiver is called background optical noise.

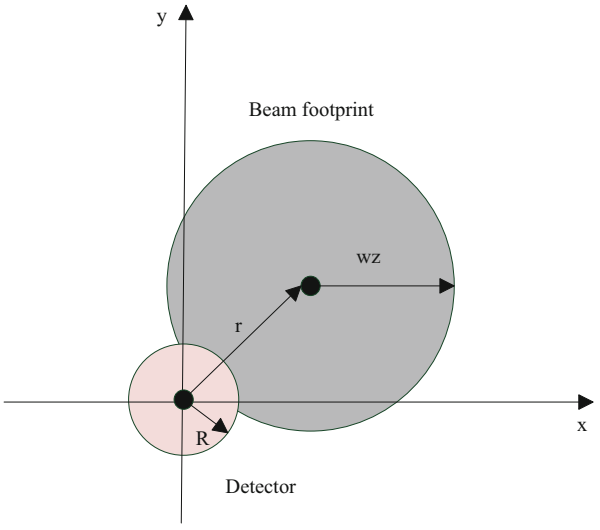


Fig. 7.2 Offset of the detector and beam on the receiving plane [47]

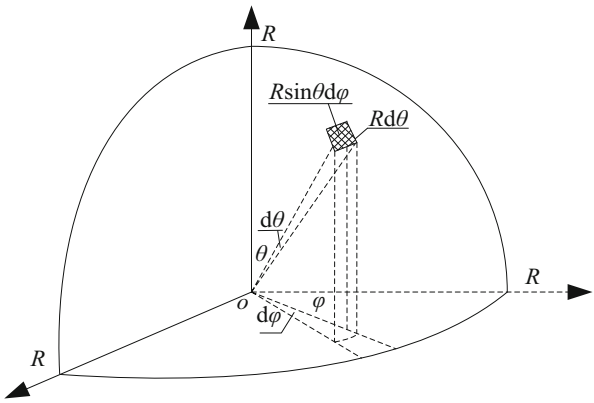


Fig. 7.3 Solid angle and bin projection in a spherical coordinate system

M_λ represents the total power radiated into the hemispherical space over a unit area and wavelength range. $L_\lambda(\theta, \varphi)$ represents the power radiated into a unit steradian angle over a unit projected area and wavelength range in a spherical coordinate system pointing in the direction of (θ, φ). The integral conversion relation between M_λ and $L_\lambda(\theta, \varphi)$ is shown in Fig. 7.3 [53] and can be expressed as

$$M_\lambda = \int_o L_\lambda(\theta, \varphi) \cos\theta \mathrm{d}\Omega, \tag{7.31}$$

where o represents the hemisphere, λ is the wavelength, θ is the angle between the beam and longitudinal axis, φ is the angle between the beam and transverse axis, and Ω is the integration region.

If the brightness of the light source is the same in each direction,

$$M_\lambda = \pi L_\lambda. \tag{7.32}$$

If the light source can be regarded as a point source,

$$M_\lambda = \frac{2\pi c^2 h}{\lambda}\left[\exp\left(\frac{hc}{k_\mathrm{B}\lambda T}\right) - 1\right]^{-1}, \tag{7.33}$$

where c is the speed of light; h is Planck's constant, value: $h = 6.6260693 \times 10^{-34}\mathrm{J} \cdot \mathrm{s}$; λ is the wavelength; k_B is Boltzmann's constant, value: $k_\mathrm{B} = 1.380649 \times 10^{-23}\mathrm{J/K}$; and T is the blackbody temperature of the radiation source.

Ke's Model

Visible-light communication (VLC) both illuminates and communicates and has attracted much attention in intelligent-transport systems. Ke et al. proposed a nighttime background-light noise model (Ke's model) using a double-Gaussian function for fitting. The expression of the noise model is [56]

$$f(I) = A_1 \times \frac{1}{\sqrt{2\pi} \times \sigma_1} \times e^{\left[-\frac{(I-\mu_1)^2}{2\times\sigma_1{}^2}\right]} + A_2 \times \frac{1}{\sqrt{2\pi} \times \sigma_2} \times e^{\left[-\frac{(I-\mu_2)^2}{2\times\sigma_2{}^2}\right]}, \tag{7.34}$$

where A_1 and A_2 are normalized correction coefficients that take values in the range of 0–1, such that $\int_0^{+\infty} f(I) = 1$; μ_1 and σ_1 are, respectively, the mean and variance of the flicker of the streetlight intensity where the vehicle is located. μ_2 and σ_2 are, respectively, the mean and variance of the flicker of the intensity of the light from other sources.

The cumulative distribution function (CDF) of the background-light noise model is fitted by a double-Gaussian function, and the expression is [58]

$$\begin{aligned} F(I) &= \int_0^I f(X)d\mathrm{X} \\ &= \int_0^I A_1 \times \frac{1}{\sqrt{2\pi} \times \sigma_1} \times e^{\left[-\frac{(X-\mu_1)^2}{2\times\sigma_1^2}\right]} + A_2 \times \frac{1}{\sqrt{2\pi} \times \sigma_2} \\ &\quad \times e^{\left[-\frac{(X-\mu_2)^2}{2\times\sigma_2^2}\right]} \mathrm{d}X, \end{aligned} \tag{7.35}$$

where X is the variable.

Gaussian curves 1 and 2 in Fig. 7.4 characterize the noise model with weak and strong light-intensity fluctuations, respectively, and the double-Gaussian model superimposed on the two Gaussian curves is the Koch model. The model well matches the actual situation and can be used to characterize the noise model of visible-light communication for vehicular networking in different regions and under different weather conditions.

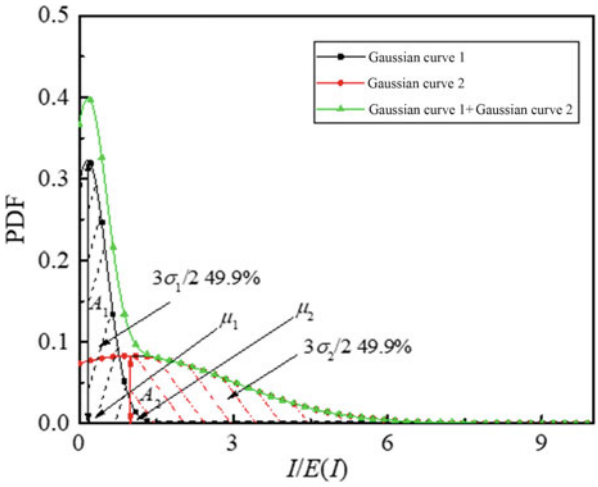


Fig. 7.4 Ke's noise model [56]

Beacon Noise

Research Progress

Beacon noise in optical wireless communication systems is mainly caused by the nonideal nature of optical receiving devices and signal processors, which can lead to signal distortion and signal-to-noise-ratio degradation. Beacon noise mainly includes thermal noise, dark-current noise, scattering noise, and 1/f low-frequency noise. In 2005, Bao et al. [65] conducted experimental and theoretical studies on the flicker noise (1/f noise) of optoelectronically coupled devices before and after electric stress and established a quantitative analysis model of 1/f noise. In 2007, Gao et al. [66] derived a mathematical model for the characteristics of avalanche photodiodes (APDs), based on Poisson's equation and the continuity equation of carriers.

In 2008, Wang [67] analytically modeled the scattering noise, thermal noise, and background noise in an optical wireless communication system and qualitatively analyzed the transmission performance of each part of the system, confidentiality performance, and other important indexes. In 2010, Hu et al. [68] elaborated on photodetector photoconversion circuits, amplifier-circuit noise, amplifier-circuit stability, etc., and designed an amplifier circuit that could effectively reduce the noise and temperature drift with a large dynamic input range.

In 2012, Wang et al. [69] theoretically analyzed the effect of background-light-induced scattering grain noise on a communication system, proposed a theoretical model that allowed for receiver sensitivity and the corresponding power loss, and experimentally verified the model. The experimental results showed that, in this system, when the operating bit rate was low, the background light led to a power loss of several decibels. The noise induced by the preamplifier increased as the bit rate increased, eventually dominating the noise process. In the same year, Wu et al. [70] explained in detail the physical model of the 1/f noise of tunnel-junction magnetoresistive sensors, built a test platform for it, and verified that the magnetic-shielding system had a good shielding effect on the ambient magnetic field.

In 2013, Kharraz et al. [71] analyzed the performance of PIN and APDs, which are commonly used photodetectors in optical communication systems, in terms of the bandwidth, gain, extinction ratio, and scattering noise. In the same year, Xu et al. [72] analyzed the time- and frequency-domain characteristics of circuit-board thermal noise and photomultiplier-tube dark-current noise in ultraviolet (UV) optical communication. In 2016, Li [73] proposed that the noise in photodetection circuits could be effectively addressed through a reasonable selection of amplifiers and amplification circuits.

In 2017, Zhai [74] illustrated the working principle and performance analysis of UV photodetectors with different structures, including photoconductive, PIN-type, Schottky-type, metal–semiconductor–metal (MSM)-type, and avalanche UV detectors. In 2019, Yan et al. [75] proposed a method to recover the clock and data directly from a discrete random-pulse sequence output from a single-photon avalanche diode. They experimentally verified the effects of the system-setup parameters,

such as the number of selective signals, length of synchronization characters, and baud rate, on the average bit-error rate.

In 2021, Xu [76] analyzed the average BER and average channel capacity of a free-space optical communication system under the influence of background noise, thermal noise, and quantum noise. In 2022, Mostafapour et al. [77] discussed the case of multiplicative noise in visible-light communication, modeled the implementation of a mobile adaptive network, and analyzed its performance. The results showed that the network had a completely different performance compared to the additive-noise case.

In the same year, Song et al. [78] investigated the effects of the receiver-temperature, average received light intensity, and magnitude of the avalanche-photodiode gain value on the reliability of an underwater optical wireless communication system subjected to the limitation of scattered grain noise. The results showed that the BER performance of the system could be improved to a certain extent by selecting a suitable avalanche-photodiode gain value. In 2023, Ke et al. [79] described a detector's scatter-grain noise, generating-complex noise, 1/f noise, and thermal noise, according to the photodetector noise-generation mechanism, and provided the corresponding mathematical model.

Noise Models

Poisson Noise Model

Poisson noise, also known as photon-counting noise or photon-jitter noise, arises primarily from the discrete nature of photons. When an optical signal passes through a detector, such as an avalanche photodiode, the randomness of the photon count causes the photocurrent to generate Poisson noise. Poisson noise is statistical noise that describes the randomness of the photon-counting process.

Poisson noise is related to the light intensity. As the fluctuation of the photon count increases with the light intensity, the Poisson noise becomes more severe. Therefore, the Poisson distribution can be used for modeling:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \tag{7.36}$$

where X represents the random variable, k is the number of occurrences per unit time, ! represents the factorial, and λ is the average occurrence rate per unit time.

When designing an optical wireless communication system, Poisson noise should be considered for its impact on the signal-to-noise ratio and bit-error rate, and corresponding measures should be taken to reduce its damage to system performance. These measures include selecting high-performance photodetectors, adopting digital signal-processing techniques, and optimizing the optical system design. The Poisson noise model has important applications in many fields, such as medical imaging, and can be used to describe the absorption of X-rays as they pass through human tissue.

Photon-Counting Noise Model

Photodiodes and avalanche diodes are often used as detection devices in optical wireless communication receiving systems. When a photon strikes the diode surface, the avalanche effect triggered by the diode generates a high current, thus counting the photons once. In the receiver, the relationship between the input optical signal S_i and the output electrical signal S_0 can be expressed as follows:

$$S_0 = S_i \cdot \mathrm{C}_P \cdot M + S_n, \tag{7.37}$$

where C_P is the photon-detection efficiency, M is the device gain, and S_n is the noise signal.

To detect the amplitude of the output impulse response, a judgment threshold can be set to ensure that the probability distribution meets the following requirements:

$$P(n, \lambda) = \frac{\lambda^n \exp(-\lambda)}{n!}, \tag{7.38}$$

where n represents the number of photoelectrons, λ represents the average pulse amplitude, and ! represents the factorial.

Multiplicative-Noise Model

In optical wireless communication systems, when using APDs and other devices with multiplication characteristics, the PN junction generates electron–hole pairs under the action of an external field. These electrons and holes move to the N and P regions, respectively, under the action of an electric field. In the high electric-field region (avalanche region), electrons and holes obtain sufficient energy to collide and dissociate to form more electron–hole pairs. This process is called avalanche multiplication.

The avalanche-multiplication process is random because it depends on the paths of the electrons and holes within the avalanche region, their velocities, and their interactions with the lattice. This randomness leads to fluctuations in the multiplication factor, resulting in avalanche multiplication noise. Avalanche multiplication noise is superimposed on the detected photocurrent signal, reducing the signal-to-noise ratio and thus affecting the performance of the communication system.

To assess the effect of avalanche multiplication noise on the system performance, the noise can be quantified by calculating the avalanche multiplication-noise factor (F_n). F_n is a function of the gain (M) of the APD and can be calculated using the following equation:

$$F_n = k \cdot M + \frac{1 - k}{M}, \tag{7.39}$$

where M is the gain of the APD. k is the ratio of the ionization coefficients of the APD and represents the relative strength of the avalanche multiplication effect produced by electrons and holes. The value of k is usually between 0 and 1 and is related to the material and structure of the APD.

When designing an optical wireless communication system, the relationship between the gain of the APD and avalanche multiplication noise must be determined. The impact of avalanche multiplication noise on the system performance can be reduced by optimizing the operating voltage of the APD, selecting a suitable gain, and using digital signal-processing techniques.

Photodetector-Amplification Noise Model

In photodetectors, the amplifier noise is an important source of noise, caused by the thermal movement of electrons across resistors and capacitors. The magnitude of this noise is related to the capacitance, resistance, and amplification of the photocurrent, and the gain of the device.

In optical receivers, in addition to the detector, other devices generate a large amount of thermal noise, particularly high-frequency preamplifiers. Because the total gain of the amplifier group is very high and the high-frequency amplification occurs first, the noise of the high-frequency amplifier will be progressively amplified in the subsequent levels. Therefore, a noise figure is required to describe the deterioration of the signal-to-noise ratio after the signal has passed through the amplifier circuit [61]. The noise factor is

$$N_F = 101g\frac{P_i}{P_o}(\text{dB}), \tag{7.40}$$

where N_F is the amplification of thermal noise by a high-frequency amplifier, P_i is the signal-to-noise power ratio at the input, P_o is the signal-to-noise power ratio at the output, and dB is the unit of the noise coefficient. It can be seen that the smaller N_F is, the lower the additional noise caused by the amplifier. Because both transistors and resistive elements in the circuit cause thermal noise, their noise factor is usually greater than 1.

Photodetector amplification noise is frequency-independent white noise. In this case, the noise can be reduced by improving the internal structure of the amplifier. Therefore, establishing and analyzing the amplified noise model of the photodetector will help optimize the structure and performance of the photodetector and improve its antinoise performance and accuracy in practical applications.

Dark-Current Noise Model

In the absence of external light of any type, there is no photocurrent output from the photodetector under ideal conditions. However, carriers excited by hot electron emission, field emission, or thermal vibration of the semiconductor lattice produce an output current called a dark current. It mainly consists of the complex current generated by carriers in the depletion layer, the diffusion current generated by minority carriers at the boundary of the depletion layer, and the leakage current at the surface.

The size of the dark current is proportional to the bias voltage and the junction area of the photodiode. When the bias voltage increases, the dark current increases with the junction area. The dark current also increases with the increase in the

operating temperature of the device. The process is essentially random fluctuation, and the resulting noise is called dark-current noise. It can be expressed as

$$\left\langle i_D^2 \right\rangle = 2q(I_D + I_L) \cdot \Delta f, \tag{7.41}$$

where I_D is the dark current and I_L is the leakage current.

Thermal Noise Model

Thermal noise, which is primarily caused by currents generated by free electrons and ions inside the material as they randomly move, can be characterized by calculating the noise-equivalent bandwidths, noise coefficients, and other relevant information.

For a single carrier, the direction of thermal motion is highly random, and the average value of its current tends to zero over time. For a large number of carriers, the number of carriers moving in one direction at one instant in time is relatively large, and the number of carriers moving in the other direction at another instant in time is also relatively large. However, they have equal chances in each direction; thus, the total current effect is zero over time.

At a given instant, the current is not always zero; however, localized charges may build transiently, giving rise to undulating currents and voltages. This thermal noise in the load resistor is further amplified using a front-end amplifier. The thermal-noise power in the load resistor R_{L} can be defined as

$$S_t(f) = \frac{2k_{\mathrm{B}}T}{R_{\mathrm{L}}}, \tag{7.42}$$

where t denotes time, f denotes frequency, T denotes absolute temperature, R_{L} denotes load resistance, and k_{B} denotes Boltzmann's constant with the value of $k_{\mathrm{B}} = 1.380649 \times 10^{-23}\mathrm{J/K}$.

It can be seen that the power spectral density of the thermal noise is temperature- and amplifier-gain-dependent and is white noise. The noise figure can be used to describe the noise level of the device.

Shot Noise Model

When light collides with the surface of the detector, a large number of electrons are generated, and the movement of the electrons at the output of the detector triggers a current pulse effect that affects the external monitor or sensor [64]. This process can be described as follows:

$$x(t) = \sum_{j=1}^{\mathrm{K(v)}} h\big(1 - Z_j\big), \tag{7.43}$$

where t represents time, v represents the volume of the detection site, j represents the j^{th} electron, $h(t)$ represents the impulse response of a single electron, Z_j represents the release time of the j^{th} electron, and K(v) is a parameter used for counting. This count

represents the number of electrons generated in a volume v with a detector area of ϖ over the entire time interval, a process known as shot noise.

1/f Noise Model

1/f noise is a class of random noise with long-range correlation, self-similarity, and nonsmoothness; it is a common rise-and-fall phenomenon in the collective motion of particles and reflects the internal characteristics of the system. 1/f noise is mainly caused by the interaction of impurities and defects with carriers in the device and can sensitively reflect the existence of various potential defects in the device.

The main characteristic of 1/f noise is that the noise power-spectral density is inversely proportional to the frequency. Specifically, the power-spectral density of 1/f current noise decreases with increasing frequency, showing a decaying trend. That is, low-frequency components contribute more to the noise, while high-frequency components contribute less to the noise.

The 1/f current-noise power-spectrum expression is [65]

$$S(f) = \frac{AI^{\beta}}{f^{\gamma}}, \tag{7.44}$$

where I is the current in the device, f is the frequency, A is determined by the structural characteristics of the device, and γ and β are specific coefficients.

The effect of 1/f noise on optical communication systems may lead to data-transmission errors and degrade the system performance. Therefore, designers of optical communication systems must take measures to minimize the effects of 1/f noise.

Performance Analysis of Optical Wireless Communication Systems Under the Influence of Turbulence

In a turbulent channel with a Malaga probability distribution, the probability-density function of I can be expressed as

$$f_I(I) = \frac{g^2 A^{(G)}}{2I} \sum_{k=1}^{\infty} {a_k}^{(G)} \left(\frac{\alpha}{\xi_g}\right)^{-\frac{\alpha+k}{2}} G_{1,3}^{3,0}\left[\frac{\alpha I}{\xi_g A_0 I_1}\sqrt{\frac{\gamma_{th}}{\gamma_0}}\middle| \begin{matrix} 1+g^2 \\ g^2, a, k \end{matrix}\right] \tag{7.45}$$

$$A^{(G)} \triangleq \frac{2\alpha^{\frac{\alpha}{2}}}{{\xi_g}^{1+\frac{\alpha}{2}}\Gamma(\alpha)} \left(\frac{\xi_g \beta}{\xi_g \beta + \Omega'}\right)^{\beta} \tag{7.46}$$

$${a_k}^{(G)} \triangleq \frac{(\beta)_{k-1}\left(\alpha\xi_g\right)^{\frac{k}{2}}}{[(k-1)!]^2} \left(\frac{\Omega'}{\xi_g + {\xi_g}^2 \beta}\right)^{k-1}, \tag{7.47}$$

where k is the wave number; $K_n(\cdot)$ is the n-order type-II modified Bessel function; α and β are parameters related to the outer scale L_0 and inner scale l_0 of turbulence, respectively; g is the ratio of the equivalent beam radius at the receiver to the standard deviation of the displacement of the alignment direction error; ξ_g is the average power of the scattered light received off-axis; Ω' is the ratio of the distribution of the average optical power related to the average optical power of the line-of-sight component, etc.

γ_{th} is the communication interruption threshold of an optical receiver, i.e., when the signal-to-noise ratio $\gamma < \gamma_{th}$, a communication system, can be considered to be interrupted. γ_0 is the average signal-to-noise ratio of the received signal at the optical receiver, and $G_{p,q}^{m,n}(\cdot|\cdot)$ represents the Meijer G-function.

Turbulence significantly affects the error rate, interruption probability, and channel capacity of optical wireless communication systems. The average error rate under the influence of turbulence can be expressed as [80]

$$\overline{R}_{\mathrm{BER}} = \int_0^{\infty} P_{\mathrm{BER}}(I) f_I(I) dI, \tag{7.48}$$

where $P_{\mathrm{BER}}(I)$ is the probability-density function of the error rate. For M-order modulation optical communication systems, the probability-density function of the bit-error rate can be expressed as

$$P_{\mathrm{BER}}(I) = 2\left(1 - \sqrt{M}^{-1}\right) erfc\left(\gamma_0 I^2\right), \tag{7.49}$$

where $erfc(\cdot)$ is the compensation error function, which can be expressed as a Meijer G-function:

$$erfc\left(\sqrt{x}\right) = \frac{1}{\sqrt{\pi}} G_{1,2}^{2,0}\left[x \middle| \begin{matrix} 1 \\ 0, \frac{1}{2} \end{matrix}\right]. \tag{7.50}$$

Substituting Eqs. (7.45) and (7.49) into Eq. (7.48), the expression for the average bit-error rate can be obtained as

$$\overline{R}_{\mathrm{BER}} = \frac{\left(1 - \sqrt{M}^{-1}\right) g^2 A^{(G)}}{\pi\sqrt{\pi} \times 2^{3-(a+k)}} \sum_{k=1}^{\infty} a_k{}^{(G)} \left(\frac{a}{\xi_g}\right)^{-\frac{a+k}{2}} \times$$
$$G_{7,4}^{2,6}\left[16\gamma_0 \left(\frac{\xi_g A_0 I_1}{\alpha}\right)^2 \middle| \begin{matrix} \frac{1-g^2}{2}, \frac{2-g^2}{2}, \frac{1-\alpha}{2}, \frac{2-\alpha}{2}, \frac{1-k}{2}, \frac{2-k}{2}, 1 \\ 0, \frac{1}{2}, -\frac{g^2}{2}, \frac{1-g^2}{2} \end{matrix}\right]. \tag{7.51}$$

Owing to the influence of turbulence, the average channel capacity of an optical wireless communication system is

$$\overline{C} = \int_0^{\infty} 1b\left(1+\gamma_0 I^2\right) f_I(I) dI \tag{7.52}$$

$$1b\left(1+\gamma_0 I^2\right) = \frac{1}{\ln 2} G_{1,2}^{1,2}\left[\gamma_0 I^2 \middle| \begin{matrix} 1,1 \\ 1,0 \end{matrix}\right]. \tag{7.53}$$

Equations (7.45) and (7.53) are substituted into Eq. (7.52) to derive the average channel capacity of the system, as follows:

$$\overline{C} = \frac{g^2 A^{(G)}}{\pi \ln 2} \sum_{k=1}^{\infty} 2^{a+k-4} {a_k}^{(G)} \left(\frac{a}{\xi_g}\right)^{-\frac{a+k}{2}} \times G_{8,4}^{1,8}\left[16\gamma_0 \left(\frac{\xi_g A_0 I_1}{\alpha}\right)^2 \middle| \begin{matrix} 1, 1, \frac{1-g^2}{2}, \frac{2-g^2}{2}, \frac{1-\alpha}{2}, \frac{2-\alpha}{2}, \frac{1-k}{2}, \frac{2-k}{2} \\ 1, -\frac{g^2}{2}, \frac{1-g^2}{2}, 0 \end{matrix}\right]. \tag{7.54}$$

The interruption probability of an optical wireless communication system can be expressed as

$$P_{\text{oug}} = P_{\text{BER}}(\gamma < \gamma_{th}) = \int_0^{\gamma_{th}} f_{\gamma}(\gamma) d\gamma. \tag{7.55}$$

The relationship between the probability-density function of the signal-to-noise ratio γ and I is

$$f_{\gamma}(\gamma) = \frac{1}{2\sqrt{\gamma\gamma_0}} f_I\left(\sqrt{\frac{\gamma}{\gamma_0}}\right). \tag{7.56}$$

Thus, it can be concluded that the average channel capacity of the system is

$$P_{\text{oug}} = \frac{g^2 A^{(G)}}{2} \sum_{k=1}^{\infty} {a_k}^{(G)} \left(\frac{a}{\xi_g}\right)^{-\frac{a+k}{2}} \times G_{2,4}^{3,1}\left[\frac{\alpha}{\xi_g A_0 I_1} \sqrt{\frac{\gamma_{th}}{\gamma_0}} \middle| \begin{matrix} 1, 1+g^2 \\ g^2, a, k, 0 \end{matrix}\right]. \tag{7.57}$$

Conclusion

Noise modeling in optical wireless communication systems is a key tool for evaluating the system performance, optimizing parameters, and designing signal-processing algorithms. Studying noise models can help investigators understand

the effect of noise on the system, predict the performance limit of the system, and provide methods to guide system design and optimization.

This chapter summarized the main research work on noise models in optical wireless communications at home and abroad and highlighted Ke's model, the turbulent noise model, thermal noise model, background-light noise model, etc. Each model has different adaptive scenarios and their characteristics, as shown in Table 7.2.

Table 7.2 Summary of noise models

Noise model	Characteristics
Laser relative-intensity noise model	Laser noise is mainly related to the average power of the beam and the bandwidth of the detector
Laser phase noise model	The laser phase-noise process can be described by the Wiener process
Turbulent noise model	When the turbulence intensity in atmospheric channels is small, the lognormal model and Γ-Γ model can be used; the K-distribution model can be used to describe the situation when the turbulence intensity in the atmospheric channel is high
Phase noise model	Phase noise can seriously affect the demodulation of subcarrier signals, leading to system-performance degradation
Aiming noise model	By designing an appropriate a/w_z ratio to make the f_{h_p} curve an increasing function, the communication system performance can be improved; otherwise, it will increase the probability of signal fading
Background-light noise model	As the receiving field-of-view angle increases, the power of the background light increases
Ke's model	This model characterizes the visible-light communication noise of vehicle-to-everything in different regions and under different weather conditions
Poisson noise model	Poisson noise is related to the light intensity; the greater the light intensity, the greater the fluctuation in the number of photons received, resulting in more severe Poisson noise
Photon-counting noise model	The photon distribution is a Poisson distribution
Multiplicative noise model	Avalanche doubling noise is random
Photodetector-amplification noise model	The magnitude of amplification noise in photodetectors is related to various factors, such as the capacitance, resistance, and the amplification factor of the photocurrent
Thermal noise model	The power-spectral density of thermal noise is related to the temperature, amplifier gain, and other factors and is a white noise. The noise figure can be used to describe the noise level of a device
Dark-current noise model	The magnitude of the dark current is proportional to the bias voltage and the junction area of the photodiode
Shot noise model	Shot noise is a strongly random noise; its size and distribution can be analyzed based on Poisson's distribution theory
1/f noise model	1/f noise has long-range correlations, and its composition is relatively complex

Current noise-modeling research still faces some challenges and shortcomings. For example, more in-depth studies are required to model background-light noise in complex environments, and more accurate modeling methods are required to determine the effects of nonlinear effects and multipath propagation on noise. In addition, more in-depth and detailed explorations are required for noise modeling in different application scenarios and system requirements.

Based on the above analyses, future research can focus on the following aspects. First, noise models can be further improved to increase their accuracy and applicability. Second, additional experimental and simulation studies should be conducted to verify the validity of the noise models and provide guidance for practical system design and optimization. In addition, new noise-modeling methods and noise-suppression strategies should be explored in combination with techniques, such as machine learning and deep learning, to improve the performance and stability of the system.

References

1. Su M S, Chen C Y. Development status of wireless optical communication in China [J]. Optoelectromechanical Information, 2004 (10): 20–23.
2. Ke X Z, Deng L J. Wireless optical communication [M]. Beijing: Science Press, 2016: 117–120.
3. Lv H, Gao Z. Lasers and detectors in Free-space optical communication system [J]. Telecom Express, 2003 (03): 35–37.
4. Yang H D, Wang R, Yang X. L.CO2 laser and its application in wireless optical communication [J]. Optical Communication Technology, 2004 (05): 52–54.
5. Gensty T, Elsäßer W. Semiclassical model for the relative intensity noise of intersubband quantum cascade lasers [J]. Optics Communications, 2005, 256(03): 171–183.
6. Yu B L, Zhen S L, Zhu J, et al. Experimental research on low noise fiber lasers [J]. Journal of Optics, 2006, 26 (02): 217–220.
7. Camatel S, Ferrero V. Narrow linewidth CW laser phase noise characterization methods for coherent transmission system applications [J]. Journal of Lightwave Technology, 2008, 26(17): 3048–3055.
8. Di Domenico G, Schilt S, Thomann P. Simple approach to the relation between laser frequency noise and laser line shape [J]. Applied optics, 2010, 49(25): 4801–4807.
9. Xie W L, Dong Y, Zhou Q, et al. Analysis of the impact of laser phase noise in optical frequency domain reflection technology [J]. Journal of Optics, 2011,31 (07): 68–73.
10. Liu J H. The Effect of Laser Phase Noise on the Performance of Coherent Fiber Optic Communication Systems [J]. Journal of Xi'an University of Posts and Telecommunications, 2012, 17 (05): 25–28.
11. Chen D, Ke X Z, Zhang L. Characteristics of laser intermodulation distortion in turbulent channel [J]. Acta Photonica Sinica, 2016, 45 (02): 99–103.
12. Warid H H. Modulation response and relative intensity noise spectra in quantum cascade lasers [J]. Physical Chemistry Research, 2017, 5(2):377–394.
13. Song Z Y, Yao G B, Zhang L L, et al. Factors affecting phase noise of single frequency fiber lasers [J]. Infrared and Laser Engineering, 2017, 46 (03): 97–100.
14. Bai Y. two μ Characterization method of m-band laser linewidth and development and application of single longitudinal mode thulium doped fiber laser [D]. Beijing: Beijing Jiaotong University, 2019.
15. Didier P, Spitz O, Cerutti L, et al. Relative intensity noise and intrinsic properties of RF mounted interband cascade laser [J]. Applied Physics Letters, 2021, 119(17): 171–177.

16. Du Y C, Zhang R, Wang L S, et al. Chaos synchronization of distributed Bragg reflection semiconductor lasers driven by common noise [J]. Journal of Optics, 2022, 42 (23): 154–161.
17. Seminara M, Gabbrielli T, Corrias N, et al. Characterization of noise regimes in mid-IR free-space optical communication based on quantum cascade lasers [J]. Optics Express, 2022, 30(25): 44640–44656.
18. Razavi M, Shapiro J H. Wireless optical communications via diversity reception and optical preamplification [J]. IEEE Transactions on wireless Communications, 2005, 4(3):975–983.
19. Kiasaleh K. Performance of APD-based, PPM free-space optical communication systems in atmospheric turbulence [J]. IEEE transactions on communications, 2005, 53(09):1455–1461.
20. Zhu J. Research on indoor wireless optical communication uplink and downlink based on white LED [D]. Jiangsu: Jiangsu University, 2008.
21. Ma B K, Guo L X. The impact of fog on the signal-to-noise ratio of wireless optical communication systems [J]. Journal of Xi'an University of Posts and Telecommunications, 2008, 67 (01): 25–27.
22. Jaruwatanadilok S. Underwater wireless optical communication channel modeling and performance evaluation using vector radiative transfer theory [J]. IEEE Journal on selected areas in communications, 2008, 26(09):1620–1627.
23. Cochenour B M, Mullen L J, Laux A E. Characterization of the Beam-Spread Function for Underwater Wireless Optical Communications Links [J].IEEE Journal of Oceanic Engineering, 2009, 33(04):513–521.
24. Meng H C. Augmented Lagrange algorithm and its application in optimization design of wireless optical communication systems [D]. Wuhan: Wuhan University of Technology, 2010.
25. Wang H X, Hu H, Zhang T Y, et al. Modeling and Simulation of Error Performance in Wireless Optical Communication Systems in Weak Turbulence [J]. Journal of System Simulation, 2011, 23 (04): 788–792.
26. Tu Y J, Zhou X L, Zhang H J, et al. Performance of free space optical communication based on mixed noise model [J]. Information and electronic engineering, 2012, 10 (06): 649–653.
27. Abd El-Naser A. Mohamed, Ahmed Nabih Zaki Rashed and Amina E. M. El-Nabawy. The Effects of the Bad Weather on the Transmission and Performance Efficiency of Optical Wireless Communication Systems [J]. International Journal of Image, Graphics and Signal Processing, 2012, 4(7):68–83.
28. Yao W M, Rao J H, Zhang X H, et al. Research on FDPIM performance in underwater wireless optical communication [J]. Laser Technology, 2013, 37 (05): 605–609.
29. Tang S, Dong Y, Zhang X. Impulse response modeling for underwater wireless optical communication links [J]. IEEE transactions on communications, 2013, 62(01):226–234.
30. Ansari I S, Yilmaz F, Alouini M S. Impact of pointing errors on the performance of mixed RF/FSO dual-hop transmission systems [J]. IEEE Wireless Communications Letters, 2013, 2(3):351–354.
31. Nistazakis H E, Stassinakis A N, Muhammad S S, et al. BER estimation for multi-hop RoFSO QAM or PSK OFDM communication systems over gamma gamma or exponentially modeled turbulence channels [J]. Optics & Laser Technology, 2014, 64: 106–112.
32. Varotsos G K, Stassinakis A N, Nistazakis H E, et al. Probability of fade estimation for FSO links with time dispersion and turbulence modeled with the gamma–gamma or the IK distribution [J]. Optik, 2014, 125(24): 7191–7197.
33. Zhu D, Zhang F, Zhou P, et al. Wideband phase noise measurement using a multifunctional microwave photonic processor [J]. IEEE Photonics Technology Letters, 2014, 26(24): 2434–2437.
34. Ke X Z, Liu M. Diversity reception technology in turbulent channel wireless optical communication [J]. Journal of Optics, 2015, 35 (01): 88–95.
35. Yang B S. Research on the symbol error rate performance of MPPM based wireless optical communication systems in turbulent channels [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2015.

36. Rashed A N Z. Optical wireless communication systems operation performance efficiency evaluation in the presence of different fog density levels and noise impact [J]. Wireless Personal Communications, 2015, 81: 427–444.
37. Xu T, Jacobsen G, Popov S, et al. Analytical BER performance in differential n-PSK coherent transmission system influenced by equalization enhanced phase noise [J]. Optics Communications, 2015, 334: 222–227.
38. Wu J P, Song L. Performance Analysis of Optical Wireless Communication System Based on K-Distributed Atmospheric Channel Model [J]. Journal of Computational and Theoretical Nanoscience, 2016, 13(1): 274–279.
39. Hamza T, Khalighi M A, Bourennane S, et al. Investigation of solar noise impact on the performance of underwater wireless optical communication links [J]. Optics Express, 2016, 24(22):25832–25845.
40. Anbarasi K, Hemanth C, Sangeetha R G. A review on channel models in free space optical communication systems [J]. Optics Laser Technology, 2017, 97:161–171.
41. Fan X K, Zhang L, Tong S F, et al. The influence of sky background light on space laser communication systems [J]. Progress in Laser and Optoelectronics, 2017, 54 (07): 102–110.
42. Zhang G F. Optimal Relay Deployment Algorithm for Free Space Optical Communication System [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2017.
43. Ke X Z, Wang C H, Chen D. Phase noise analysis of subcarrier modulation system in Malaga atmospheric turbulence channel [J]. Journal of Communications, 2018, 39 (11): 80–86.
44. Wang C H. Research on Phase Noise Characteristics and Compensation Techniques of Wireless Optical Subcarrier Modulation [D]. Xi'an: Xi'an University of Technology, 2019.
45. Xing F Y. Research on Several Key Technologies of Underwater Wireless Optical Network Collaborative Communication [D]. Dalian: Dalian University of Technology, 2020.
46. Hong Z, Yan Q, Li Z, et al. Monte Carlo Simulation and Implementation of Underwater Single Photon Communication System [J]. Second Target Recognition and Artificial Intelligence Summit Forum, 2020, 11427(34):792–798.
47. Fu Y Q, Duan Q, Zhou L. Performance Study of Underwater Wireless Optical Communication System with Gamma Gamma Strong Ocean Turbulence and Aiming Error [J]. Infrared and Laser Engineering, 2020, 49 (02): 110–117.
48. Wang X, Su X, Liu G, et al. A method for improving the detection accuracy of the spot position of the four-quadrant detector in a free space optical communication system [J]. Sensors, 2020, 20(24): 7164–7165.
49. Liu W J. Research on Non ideal Transmission Characteristics of Underwater LED Wireless Optical Communication System [D]. Xi'an: Xi'an University of Posts and Telecommunications, 2020.
50. Wang L J. Research on static node self-organizing network clustering algorithm based on UWOC [D]. Chengdu: University of Electronic Science and Technology, 2020.
51. Zhang L, Zhou X, Du J, et al. Fast self-learning modulation recognition method for smart underwater optical communication systems [J]. Optics Express, 2020, 28(25): 38223–38240.
52. Ke X Z, Wu J L, Yang S J. Research progress and prospects of atmospheric turbulence for wireless optical communication [J]. Journal of Radio Science, 2021, 36 (03): 323–339.
53. Lee-Leon A, Yuen C, Herremans D. Underwater acoustic communication receiver using deep belief network [J]. IEEE Transactions on Communications, 2021, 69(6): 3698–3708.
54. Xu Z X, Xu G J, Zheng Z Q. BER and Channel Capacity Performance of an FSO Communication System over Atmospheric Turbulence with Different Types of Noise [J]. Sensors, 2021, 21(10):3454–3454.
55. Han Z D. Research on Visible Light Indoor Positioning System Based on Neural Networks [D]. Xi'an: Xi'an University of Technology, 2021.
56. Ke X Z, Qin H H, Yang S J, et al. A Night Background Light Noise Model for the Visible Light Communication System of the Internet of Vehicles [J]. Journal of Radio Science, 2021, 36 (06): 986–990.
57. Majlesein B, Gholami A, Ghassemlooy Z. Investigation of the scattering noise in underwater optical wireless communications [J]. Sci, 2021, 3(2): 27.

58. Zhang J L, Gao G J, Wang J J. Theoretical and experimental verification of underwater wireless optical communication under noisy light [J]. Journal of Beijing University of Posts and Telecommunications, 2022, 45 (03): 102–106.
59. Li Z, Liao Z W, Liang J Y, et al. Research and Prospects of Atmospheric Turbulence Models [J]. Optical Communication Technology, 2022, 15 (03): 1–13.
60. Chen E, Buscaino B, Kahn J M. Phase Noise Analysis of Resonator-Enhanced Electro-Optic Comb-Based Analog Coherent Receivers [J]. Journal of Lightwave Technology, 2022, 40(21): 7117–7128.
61. Yang Y, Lei J Y, Wang X B, et al. Research on the performance of the composite seawater channel offset four phase shift keying modulation system [J]. Acta Photonica Sinica, 2022,51 (10): 203–215.
62. Ke X Z, Liao Z W, Liang J Y, et al. Measurement of atmospheric turbulence noise in wireless optical communication [J]. Journal of Radio Science, 2023, 6 (4): 1–13.
63. Jin C, Shevchenko N A, Wang J, et al. Wideband Multichannel Nyquist-Spaced Long-Haul Optical Transmission Influenced by Enhanced Equalization Phase Noise [J]. Sensors, 2023, 23(3): 1493–1493.
64. Li Z, Wang F G, Liang J Y, et al. Research progress in LED visible light communication [J]. Journal of Lighting Engineering, 2023, 34 (01): 29–44.
65. Bao J L, Zhuang Y Q, Du L, et al. Flicker noise model of optocoupler [J]. Acta Photonica Sinica, 2005, 34 (09): 1359–1362.
66. Gao X J, Zhang X C, Chen Y. Device Model and Numerical Simulation of InGaAs/InP SAGCM-APD [J]. Semiconductor Optoelectronics, 2007, 28 (05): 617–622.
67. Wang X Y. Design of Optical Receiver for Image Wireless Optical Communication System [D]. Changchun: Changchun University of Science and Technology, 2008.
68. Hu T, Si H Y. Design and Research of Preamplifier Circuit for Photoelectric Detectors [J]. Application of Optoelectronic Technology, 2010, 25 (01): 52–55.
69. Wang K, Nirmalathas A, Lim C, et al. Impact of background light induced shot noise in high-speed full-duplex indoor optical wireless communication systems [J]. Optics express, 2011, 19(22): 21321–21332.
70. Wu S B, Chen S, Li H, et al. Research progress on 1/f noise of TMR and GMR sensors [J]. Journal of Physics, 2012, 61 (09): 552–561.
71. Kharraz O, Forsyth D. Performance comparisons between PIN and APD photodetectors for use in optical communication systems [J]. Optik, 2013, 124(13): 1493–1498.
72. Xu H R, Zuo Y, Zhang W B, et al. Research on Noise in Ultraviolet Optical Communication Systems [J]. Optical Communication Research, 2013 (02): 65–69.
73. Li R L. Research on Noise Analysis and Noise Treatment of Optoelectronic Detection Circuits [J]. Industry and Technology Forum, 2016,15 (13): 43–44.
74. Zhai Y. Study on photoelectric characteristics of GaN based ultraviolet detectors [D]. Wuxi, Jiangsu: Jiangnan University, 2017.
75. Yan Q, Li Z, Hong Z, et al. Photon-counting underwater wireless optical communication by recovering clock and data from discrete single photon pulses [J]. IEEE Photonics Journal, 2019, 11(05):1–15.
76. Xu Z, Xu G, Zheng Z. BER and channel capacity performance of an FSO communication system over atmospheric turbulence with different types of noise [J]. Sensors, 2021, 21(10): 34–54.
77. Mostafapour E, Ghobadi C, Nourinia J, et al. Performance analysis of mobile adaptive networks in VLC multiplicative SPAD channel noise conditions [J]. Optics Communications, 2022, 5(24): 128760–128763.
78. Song Y T, Li S, Li G, et al. Performance of partially coherent optical OFDM-UWOC system considering APD [J]. Journal of Xidian University, 2023,50 (01): 58–65.
79. Ke C H, Bao J, Liang J Y, et al. Research progress on detectors and their noise models [J]. Optical Technology, 2023,49 (03): 329–348
80. Ji X Y, Yin H X, Jing L Y, et al. Performance analysis of underwater wireless optical communication systems in strong turbulent channels based on strong wave theory [J]. Journal of Optics, 2022, 42 (18): 9–16.

Adaptive-Threshold Detection Technology

8

Contents

Introduction 238
Domestic and Foreign Research Progress 239
Adaptive-Threshold Detection 239
Foreign Research Progress 240
Domestic Research Progress 242
System Structure and Model 245
Optical Wireless Communication System Structure 245
Photoelectric Detection Noise 245
Error Rate 246
Optimal Threshold Detection 247
Typical Adaptive-Threshold Detection Models 250
LMS Filter Adaptive-Threshold Detection Model 250
Kalman-Filter Adaptive-Threshold Detection Model 251
Adaptive-Threshold Detection Model with a Fading Kalman Filter 253
High-Order Cumulant Adaptive-Threshold Detection 254
Decision-Guided Adaptive-Filtering Algorithm 254
Sliding-Correlation Adaptive-Threshold Algorithm 255
Adaptive-Threshold Detection in a Strong-Turbulence Channel 257
Summary 263
References 264

Abstract

The transmission of optical signals in atmospheric turbulent channels causes signal fading and light intensity flicker, and the baseband signal cannot be recovered correctly by fixed threshold detection at the receiving end, so adaptive adjustment of the received signal decision threshold is required. Adaptive threshold detection technology can effectively suppress the atmospheric turbulence effect, which is an important means to improve the bit error rate performance of optical wireless communication systems and enhance system reliability. Its detection performance is mainly optimized and

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_8

improved for the threshold detection algorithm and feedback mechanism. Reviewing the development process of adaptive threshold detection technology, starting from the structure of optical wireless communication system, deriving the optimal decision threshold model for the received signal based on Bayesian maximum likelihood estimation and maximum posterior probability criteria, and the realization of baseband signal demodulation by comparing the received signal with the optimal decision threshold. The typical adaptive threshold detection models based on the minimum mean square error filter, Kalman filter, and fading Kalman filter are analyzed, which are suitable for stationary input signals and nonstationary input signals, respectively. At the same time, the related research work of Xi′an University of Technology using high-order cumulants instead of traditional second-order statistics in the field of adaptive threshold detection is introduced. Finally, the future development trend of this field is summarized and foreseen, which can provide some reference for the future research and development of adaptive threshold detection technology for optical wireless communication.

Keywords

Optical wireless communication · Atmospheric turbulence · Adaptive threshold detection

Introduction

When an optical signal is transmitted in a turbulent atmospheric channel, the signal fades and the light intensity flickers, and the receiver cannot correctly recover the baseband signal using fixed-threshold detection [1]. Therefore, it is necessary to adaptively adjust the decision threshold of the received signal. This chapter reviews the development of adaptive-threshold detection technology.

Starting from the structure of optical wireless communication systems, based on Bayesian maximum-likelihood estimation and the maximum a posteriori probability criterion, the optimal decision threshold model of the received signal is derived, and the baseband signal is demodulated by comparing the received signal with the optimal decision threshold.

This chapter analyzes typical adaptive-threshold detection models based on three algorithms: least mean-squared error filtering, Kalman filtering, and fading Kalman filtering, which are applicable to stationary and nonstationary input signals. In addition, it introduces the research work of Xi'an University of Technology in the field of adaptive-threshold detection and finally summarizes and projects the development trends of this field in the future. It can provide a reference for the research and development of adaptive-threshold detection technology in future optical wireless communications [2, 3].

Domestic and Foreign Research Progress

Adaptive-Threshold Detection

In a optical wireless communication system, the received signal must be compared with the threshold value after receiving an optical signal. When the signal is greater than the threshold, the output is "1," and when the signal is lower than the threshold, the output is "0." Affected by atmospheric scintillation, the amplitude of the received signal fluctuates randomly, as shown in Fig. 8.1.

Errors can easily occur when using a fixed threshold. Therefore, it is necessary to study adaptive thresholds. Adaptive-threshold detection automatically adjusts the decision threshold according to changes in the received signal strength. Compared with fixed-threshold detection, adaptive-threshold detection has a wide dynamic range, and fast establishment and tracking speeds.

Using binary-threshold detection [4] as an example, adaptive-threshold detection is shown in Fig. 8.2. After receiving a signal, the receiver compares the sample value $r\ (k-1)$ of the current bit signal with the optimal decision threshold $U_T(k-1)$ and obtains the decision result, according to the decision criteria. If $r\ (k-1)$ is greater than $U_T\ (k-1)$, the decision result outputs a "1"; otherwise it outputs a "0." Then, the mean and variance of the current bit signal are estimated according to a certain amount of sampling data, and input into the adaptive predictor, which predicts the mean and variance of the signal the next time through the prediction algorithm. Then, the optimal decision-threshold value $U_T\ (k)$ at the next time is obtained and the

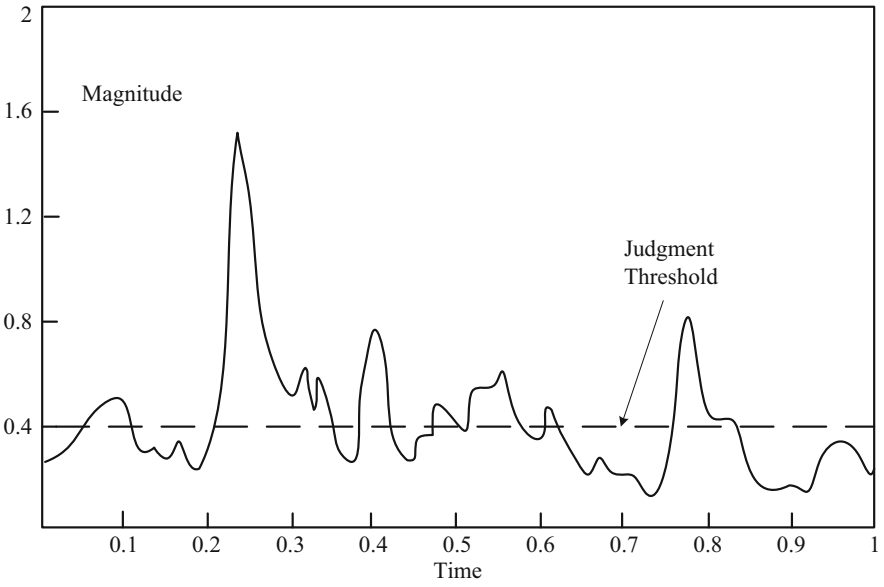


Fig. 8.1 Signal threshold detection waveform

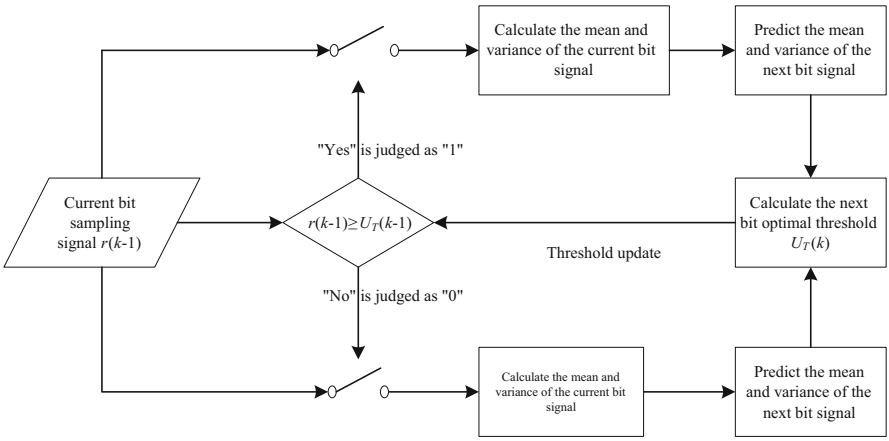


Fig. 8.2 Schematic diagram of adaptive-threshold detection [4]

threshold is updated, so that $U_T(k)$ can make a comparative decision about the next bit signal.

Foreign Research Progress

Threshold detection technology [5] appeared in the 1960s and was first proposed by Urkowitz, based on signal energy, as an unknown signal-detection method. Because this method is greatly affected by noise and has certain limitations, it has been continuously studied and improved since then. A fixed threshold has been used since early signal detection. With the increasing communication-quality requirement, an adaptive threshold, which automatically adjusted the decision threshold with the change in received signal strength, was urgently needed [6]. Therefore, adaptive-threshold detection was proposed in the late 1990s.

In 2001, Mukai et al. [7] used adaptive threshold detection technique in deep space optical communication, which uses the average signal strength to estimate the degree of signal fading and total attenuation in response to changing channel conditions. However, the adaptive threshold detection technique used is not yet mature and requires further research.

In 2002, Burris et al. [8] introduced Kalman estimation for adaptive-threshold detection in optical wireless communication. In an optical communication receiver with multiplicative noise, the likelihood-ratio test (LRT) scheme was used. The Kalman filter algorithm was used to track the mean and variance of the bits and update the detection threshold.

In 2003, Kahn et al. [9] investigated the effect of three time-domain detection techniques, Maximum-Likelihood (ML) symbol-by-symbol detection, Maximum-Likelihood Sequence Detection (MLSD), and Pilot-Symbol (MLSD) and Pilot-Symbol Assisted Detection (PSAD) techniques to mitigate the light intensity

fluctuations caused by atmospheric turbulence. Experimental results show that MLSD and PSAD obtain signal-to-noise gains of 2.4 dB and 1.9 dB, respectively, compared to ML-by-symbol. Since MLSD exploits the temporal correlation of intensity fluctuations, its detection performance is better than the other two time-domain detection techniques and thus has been widely studied since then, but its cost and complexity are high.

In 2004, Cole et al. [10] proposed using the minimum mean-square error (MMSE) to estimate the fading degree of a signal. In weak turbulence, the detection performance of this algorithm was significantly higher than that of the maximum-likelihood algorithm.

In 2005, Namazi et al. [11] proposed a new method for the binary-data synchronization and detection of free-space optical signals, based on wavelet transform and adaptive Wiener filtering. The fixed-scale Haar wavelet transform was analyzed, the bit-error probability was obtained using the Bayesian ternary hypothesis test, and the optimal threshold detection model was deduced.

In 2007, Schober et al. [12] investigated Decision-Feedback Detection (DFD) link for symbol-by-symbol decision using the observation window of the previous decision and reception statistics. Simulation results show that the performance of the judgment-feedback link approaches the lower limit of coherent detection as the window length increases.

In 2009, Rieger et al. [13] proposed an optimal multisymbol detection (MSD) fast-search algorithm for incoherent detection in space optical communication, as well as suboptimal multisymbol detection with a low-complexity decision metric. This significantly improved the detection performance of optical wireless communication with long sequences; however, the algorithm required extensive calculation time.

In 2010, Moradi et al. [14] further studied pilot-symbol-assisted detection technology based on Kahn et al. When the channel coefficients were time-dependent, pilot-symbol-assisted detection was used to modify the maximum-likelihood function and could be extended to multiple-input multiple-output optical communication systems. However, this detection technology required a time-joint and edge probability-distribution density function, which increased the complexity of the receiver.

In 2010, Chatzidiamantis et al. [15] proposed a low-complexity suboptimal detection scheme based on generalized MLSD, considering the high computational complexity of optimal MLSD in practical applications. The scheme did not require any channel information and only used the observations of the received sequence, avoiding a complex decision process; however, it had poor performance in small observation windows.

In 2013, Kam et al. [16] designed a generalized maximum-likelihood ratio test receiver, based on maximum-likelihood sequence detection and the generalized likelihood ratio test. It did not require channel-state or channel-model information. It used the data symbols detected in the past to estimate the instantaneous channel gain and transmit power and adaptively adjusted the decision threshold. Compared

with a receiver that needs channel-model information, it has better robustness but needs a pilot at the beginning of the data transmission.

In 2013, Beaulieu et al. [17] deduced the optimal signal-detection formula for Laplace white noise and compared it with a digital matched filter for Gaussian noise. The simulation results showed that the signal-to-noise ratio (SNR) of this method was twice that of a Gaussian-noise matched filter.

In 2017, Dabiri et al. [18] considered the complexity and cost of MLSD and differential signal technology and adopted a blind estimation algorithm for on–off keying (OOK) signals. It achieved a lower bit-error performance without increasing the complexity of the receiver; however, the data length in the observation window needed to be adjusted frequently.

In 2019, Choi et al. [19] inserted single-tone guide frequency in the transmit signal to estimate the turbulent channel information for satellite optical communication, and the extracted channel state information is close to the full channel state information, which can mitigate the scintillation effect caused by turbulent atmosphere, but complex filtering process is required to suppress the crosstalk.

In 2019, Shin et al. [20] proposed an OOK differential-detection technology to solve the problem in which the implementation complexity of adaptive-threshold detection increased with an increase in the data transmission rate. It used the best decision threshold with a sequence-symbol differential method to reduce the implementation complexity.

In 2021, Choi et al. [21] proposed an adaptive-threshold decision mechanism using a random-approximation algorithm to predict the next-frame decision threshold for multilevel pulse amplitude modulation in optical wireless communication. This technology adaptively tracked the dynamics of intensity fluctuations, smoothed the intensity fluctuations, including noise signals, and improved the performance of the receiver.

Domestic Research Progress

In 2006, Wang et al. [22] used the adaptive-Kalman and homomorphic filters of a linear state model to predict the signal and turbulence statistical characteristics contained in the detection threshold. It used adaptive-threshold detection under different signal-to-noise ratio conditions and had a good inhibitory effect on atmospheric turbulence noise. The simulation and experimental results showed that under weak turbulence, the bit-error rate (BER) of the communication system with an SNR of 10–20 dB was better than that of the system with adaptive-threshold detection technology.

In 2008, Wu et al. [23] studied an adaptive-threshold detection method for signal and noise power and proposed an adaptive-threshold detection method for comparison and decisions in a sliding window. This method effectively eliminated false alarms when the signal was small and captured the signal more accurately.

In 2008, Wang [24] proposed a decision-directed (DD) algorithm based on the third-order cumulant. It used the higher-order cumulant as the cost function to update the tap coefficients of the adaptive filter, which effectively suppressed additive noise. They also verified the insensitivity of the higher-order cumulant to Gaussian noise.

In 2009, Zhao [25] proposed an adaptive sliding correlation detection method that performed cumulative correlation on the received signal, according to the sliding correlation-detection principle, to obtain the adaptive decision threshold. This method reduced the false-alarm probability and improved the probability of detecting the received signal, without increasing the system complexity.

In 2012, Wang et al. [26] studied the effect of a dynamic decision threshold (DDT) scheme adopted by the receiver to improve the bit-error performance of a coherent-optical communication system. The scheme detected the received optical signal frame by frame and provided a decision threshold for signal detection by monitoring the preamble in each detection frame, which is only applicable to low-data-rate communication. The experimental results showed that when the flicker index was 1.0, the dynamic decision threshold scheme achieved a 10^{-3} bit-error rate.

In 2013, Li et al. [27] fitted an optimal synchronization-threshold curve by analyzing the change rule of the synchronization threshold under different signal-to-noise ratios and proposed an optimal threshold-selection method. The system performance of this method was close to that of optimal threshold detection under an additive Gaussian white-noise channel, which was approximately 3 dB better than the fixed-threshold system performance.

In 2014, Shang et al. [28] proposed an improved algorithm to adjust the threshold, based on the signal-to-noise ratio according to the parameter setting. It was intended to address the low detection probability of the adaptive-threshold decision pseudo-random (PN) code-acquisition method for low signal-to-noise ratios. Monte Carlo simulation results showed that the algorithm had a higher detection probability, lower false-alarm probability, better anti-interference ability, and strong practicability.

In 2015, Zhang et al. [29] proposed an adaptive-threshold detection algorithm for the variable threshold factor estimation of noise energy using the k^{th} small element to replace the ranking noise-energy estimation. This algorithm only used a few iterations in the sorting algorithm, which greatly reduced its computational load and enabled fast adaptive-threshold decision detection without reducing the detection performance.

In 2015, Tian [30, 31] studied the problem of signal detection in strong atmospheric-turbulence channels. The experimental results showed that adaptive-threshold detection could significantly reduce the bit-error rate of optical communication systems in K-distributed turbulent channels. The more symbol samples, the better the system performance, but the higher the algorithm complexity.

In 2017, Dang et al. [32] studied an adaptive-threshold decision scheme for optical wireless communication systems without channel-state information. Based on the low-pass characteristics of atmospheric turbulence, a low-pass filter extracted instantaneous intensity fluctuations from the optical carrier component; the filtered output could be directly used as the decision threshold. Monte Carlo simulation results showed that the performance of this scheme was very close to the lower decision limit of ideal channel-state information for high-speed data transmissions.

In 2018, Zhang et al. [33] introduced a fading-Kalman filter algorithm into the adaptive-threshold detection process of optical wireless communication, effectively suppressing the impact of atmospheric turbulence. The simulation results showed that, when SNR = 10, the error rate of the adaptive-threshold detection using the

fading Kalman filter algorithm was approximately four orders of magnitude higher than that of the fixed threshold.

In 2018, Zhang et al. [34] studied the bit-error rate performance of duplex reverse-modulation optical wireless communication systems with adaptive decision thresholds under weak and strong turbulence conditions, considering the influence of detector noise. While solving for the optimal adaptive decision threshold and analyzing the bit-error rate performance, an expression for calculating the optimal adaptive decision threshold of the duplex reverse-modulation optical-communication system was also derived. Simulation results showed that an adaptive decision threshold could improve the bit-error performance of an optical communication system.

In 2018, Zhu et al. [35] proposed an adaptive-threshold detection method, based on the quantized voltage of the determined signal, to obtain the current symbol decision threshold. The DC component of the signal was extracted as a continuous-wave reference light to estimate the fluctuation characteristics of the atmospheric-turbulence channel. This reduced the high cost caused by the reference light carrying no information, and the power loss was lower.

In 2018, Geng et al. [36] designed two adaptive thresholds at the receiving end to detect the same and orthogonal branches of the signal; they did not require instantaneous channel-state information or a probability-density function. To improve the accuracy of the threshold, a decision-aid method for setting the second detection threshold was also provided. The simulation results showed that the performance of the double adaptive-threshold detection system was close to that of an ideal adaptive detection system.

In 2019, Li et al. [37] studied the error-rate performance under atmospheric-turbulence fading with a large-range correlation and proposed a new blind-detection method based on an adaptive threshold. The closed expressions of the instantaneous and average bit-error rates using the optimal adaptive threshold were derived and provided some references for the design of optical wireless communication systems.

In 2019, Jiang et al. [38] designed a new signal-detection algorithm that used stochastic-resonance technology to transform part of the noise energy into useful signal energy to improve the system's signal-to-noise ratio. It then used a two-threshold two-step algorithm for signal detection. The algorithm improved the accuracy of detection signals and reduced the interference between signals.

In 2020, Zhang et al. [39] proposed a local adaptive-threshold frequency-hopping (FH) signal-extraction and detection method. Based on a short-time Fourier transform, the appropriate local adaptive threshold was solved to eliminate the fixed-frequency interference, extract the frequency-hopping signal in the communication, and detect the extraction result. Simulation results showed that this method could effectively detect FH signals and had low complexity and high reliability.

In 2020, Liu [40] proposed a 16-way quality amplitude modulation (16-QAM) carrier-lock detection algorithm based on an adaptive threshold. Through a statistical analysis of input signals, the threshold was adjusted in real time, according to the input-signal distribution. The proposed adaptive threshold better adapted to carrier-

locking detection under different data conditions and improved the detection reliability. This was verified in satellite-docking experiments.

In 2022, Liu et al. [41] proposed a local adaptive-threshold method that was calculated in a small local window. It effectively reduced the probability of false and missed alarms and had good background-suppression ability and high detection accuracy.

System Structure and Model

Optical Wireless Communication System Structure

The structure of a optical wireless communication system [42, 43] is shown in Fig. 3.1. It mainly includes an optical transmitter, atmospheric channel, and optical receiver. The optical transmitter is composed of a signal source, modulator, laser, driving circuit, and transmitting telescope. The transmitter's function is to convert the generated electrical signal into an optical signal by modulating it onto the optical carrier, and then transmit it using the optical transmitting antenna.

The beam emitted by the transmitting antenna reaches the receiving end after transmission through a turbulent atmospheric channel. The optical receiver is composed of an optical receiving antenna, optical filter, photoelectric detector, and decision apparatus. Its function is to filter the received optical signal, send it to a photoelectric detector for a photoelectric conversion and comparison, demodulate it into an electrical signal, and recover the transmitted data.

Photoelectric Detection Noise

In long-distance optical wireless communication systems, photoelectric detectors are required to have high reception sensitivity, and avalanche photodiode (APD) detectors are generally used. The APD is a P-N structure optical-detection diode that uses the avalanche effect of carriers to amplify the photoelectric signal. However, the detector output current contains thermal, shot, amplifier, undercurrent, and other noise sources, which affect the optimal decision threshold of the system.

Thermal and shot noise are the main noise sources of photoelectric detectors. They are independent of each other and can be approximated as zero-mean Gaussian noise. Thermal noise [44] can be expressed as the mean-square thermal-noise current:

$$\sigma_T^2 = 4kTfG^2/R_L, \tag{8.1}$$

where k is the Boltzmann constant, T is the thermodynamic temperature, F is the working bandwidth of the detector, G is the amplifier gain, and R_L is the load resistance.

Shot noise describes the random fluctuation in the photoelectron-generation process, and its variance is related to the signal strength. It is expressed as the mean-square shot-noise current [45]:

$$\sigma_S^2 = 2efMNG\langle U_S\rangle = 2efM^2NG^2\langle P_R\rangle R, \tag{8.2}$$

where the signal voltage $U_S = MGP_RR$, P_R is the optical receiving power, $\langle\rangle$ represents the ensemble average, and e is the electronic charge. M is the APD multiplication factor, N is the APD noise coefficient, and R is the detector response when $M = 1$.

Error Rate

In an optical wireless communication system, the bit-error rate refers to the probability that the received symbol is misjudged. This is an important performance index for evaluating communication systems. It is composed of the false-alarm probability (P_F) and missing-alarm probability (P_M).

A false-alarm probability refers to the probability of misjudging a transmitted "0" as "1" when the amplitude of the source noise voltage exceeds the detection threshold. The probability of a missing alarm refers to the probability that the sum of the noise and signal voltages is less than the detection threshold, and a "1" is misjudged as a "0" when transmitting.

The false-alarm probability [46] can be expressed as

$$P_F = \frac{1}{2}erfc\left(\frac{U_T - U_0}{\sqrt{2\sigma_0^2}}\right). \tag{8.3}$$

The probability of a missing alarm [46] can be expressed as

$$P_M = \frac{1}{2}erfc\left(\frac{U_1 - U_T}{\sqrt{2\sigma_1^2}}\right), \tag{8.4}$$

where $erfc()$ represents the complementary error function, U_T represents the decision threshold, $\langle U_0\rangle$ and σ_0^2 represent the mean and variance of the signal voltage at "0," respectively, and $\langle U_1\rangle$ and σ_1^2 represent the mean and variance of the signal voltage at "1," respectively.

If "0" and "1" are sent, the error rate [46] of the system is expressed as

$$BER = \frac{1}{4}erfc\left(\frac{U_T - \langle U_0\rangle}{\sqrt{2\sigma_o^2}}\right) + \frac{1}{4}erfc\left(\frac{\langle U_1\rangle - U_T}{\sqrt{2\sigma_1^2}}\right). \tag{8.5}$$

Optimal Threshold Detection

To recover the information of a transmitted signal, the receiver must set a critical value for comparison with the received signal and demodulate the signal. The critical value is the threshold value. A value higher or lower than the threshold causes different changes in the signal-decision state.

Taking the binary return-to-zero code as an example, the decision threshold is typically set to 0.5. If the received-signal sampling value is greater than 0.5, it will be judged as a "1"; otherwise, it will be judged as a "0." Generally, an optimal decision threshold that can minimize the bit-error rate of the system can be found to enable the optimal reception of the received signal by the receiver.

In an adaptive receiver, optimal threshold detection [47] is based on the Bayesian likelihood-ratio test (LRT), according to the maximum a posteriori (MAP) criterion. It involves estimating and predicting the mean and variance of the signal level and relevant parameters of the atmospheric channel to obtain an optimal decision threshold and achieve adaptive reception. Therefore, the optimal decision threshold corresponding to the minimum bit-error rate is not a constant value; rather, it is a variable value that changes with atmospheric turbulence.

(1) Receiving-Signal Analysis

The receiver's adaptive-threshold decision system is shown in Fig. 8.3. It is composed of an optical receiver, front-end symbol estimator, decision circuit, and adaptive-threshold estimator and predictor [48]. The receiver first uses maximum-likelihood sequence estimation to estimate the front-end symbols, predicts and tracks the received signal, and then predicts the changes in the mean and variance of the symbols. It then obtains the optimal detection-threshold value to ensure that the bit-error rate is close to the theoretical minimum value. It compares the received signal with the obtained optimal detection threshold, makes a decision in the decision circuit, and demodulates the baseband signal.

Assuming that the received signal follows a lognormal distribution and adopts OOK modulation, the received signal model [49] can be expressed as

$$r(t) = \frac{\gamma e\lambda}{hc} I_0(t) \exp[2\chi - 2E(\chi)] + n_1(t) + n_2(t), \tag{8.6}$$

where γ is the quantum efficiency of the detector, λ is the laser wavelength, h is the Planck constant, c is the propagation speed of light in a vacuum, and $\frac{\gamma e\lambda}{hc}$ is defined as

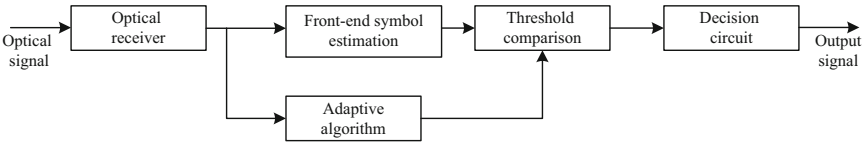


Fig. 8.3 Adaptive-threshold decision system [49]

the photoelectric-conversion efficiency η. $I_0(t)$ is the light intensity of the received signal without considering atmospheric turbulence, χ is the logarithmic amplitude fluctuation, $E[\chi]$ is the mean value of χ, $n_l(t)$ is the sum of Gaussian noise, such as thermal noise and amplifier noise unrelated to the received signal, $n_2(t)$ is the multiplicative noise related to the received signal, and $n_1(t)$ and $n_2(t)$ are statistically independent.

(2) **Maximum-Likelihood Estimation**

To achieve the best reception effect, the receiver generally uses maximum-likelihood estimation (MLE) to estimate the actual input value through several observations without any prior knowledge [50].

Assuming that the received data is optimally sampled and the symbols have been synchronized, the conditional-probability distribution function [51] when the transmitter sends a "0" (H0) is expressed as

$$p(r|H_0) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{r^2}{2\sigma_n^2}\right), \tag{8.7}$$

where σ_n^2 is the detector noise variance; $\sigma_n^2 = \sigma_{n_1}^2 + \sigma_{n_2}^2$; $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$ are the variances of noises $n_1(t)$ and $n_2(t)$, respectively; and r is the received signal.

The conditional-probability distribution function [51] when sending a "1" (H1) is

$$p(r|H_1) = \frac{1}{\sqrt{2\pi\sigma_n^2}} f_\chi(\chi) \exp\left(-\frac{\left(r - \frac{\gamma e\lambda}{hc} I_0(t) \exp[2\chi - 2E(\chi)]\right)^2}{2\sigma_n^2}\right), \tag{8.8}$$

where $f_\chi(\chi)$ is the probability density function of the logarithmic amplitude fluctuation χ.

Assuming that the probability of sending a "0" is equal to the probability of sending a "1," that is, $p(H_0) = p(H_1) = 1/2$, according to the maximum posterior probability criterion and the Bayesian formula, the maximum-likelihood detection output bit $\hat{s}$ [52] per symbol can be deduced as follows:

$$\hat{s}_{MLE}(r) = \underset{s}{^{argmax}} p(r|s). \tag{8.9}$$

The likelihood ratio (LR) is obtained from the above derivation [53]:

$$\wedge(r) = \int_{-\infty}^{\infty} f_\chi(\chi) \exp\left[-\frac{\left(r - \frac{\gamma e\lambda}{hc} I_0(t) \exp[2\chi - 2E(\chi)]\right)^2 - r^2}{2\sigma_n^2}\right] \begin{matrix} > H_1 \\ < H_0 \end{matrix} 1, \tag{8.10}$$

where $\wedge(r)$ is the likelihood ratio of the received signal, $f_\chi(\chi)$ is the probability-density function of the logarithmic amplitude fluctuation, σ_n^2 is the detector noise

variance, and I_0 (t) is the light intensity of the received signal without considering atmospheric turbulence.

(3) **Optimal Decision Threshold**

To improve the system-detection performance, the false-alarm probability P_F and missed-alarm probability P_M must be as small as possible. The detection threshold satisfying this condition is called the optimal decision threshold. The optimal decision threshold U_{OT} can be expressed as [54]

$$U_{OT} = \text{argmin}(P_F + P_M), \tag{8.11}$$

where argmin() represents the minimum value of the objective function, P_F represents the false-alarm probability, and P_M represents the missed-alarm probability. Equation (8.11) can be solved by finding the zero point using a partial differential [54], that is,

$$\frac{\partial P_F}{\partial U_T} - \frac{\partial P_M}{\partial U_T} = 0. \tag{8.12}$$

Combining Eqs. (8.3) and (8.4), Eq. (8.12) can be converted into

$$-\frac{1}{2} erfc\left(\frac{U_T - U_0}{\sqrt{2\sigma_0^2}}\right) = -\frac{1}{2} erfc\left(\frac{U_1 - U_T}{\sqrt{2\sigma_1^2}}\right), \tag{8.13}$$

where U_0, U_1 and σ_0^2, σ_1^2, respectively, represent the mean and variance of the signal voltage at "0" and "1." Taking the natural logarithm of both sides of Eq. (8.13) and sorting it out:

$$\left(\sigma_1^2 - \sigma_0^2\right)U_T^2 + \left(2U_1\sigma_0^2 - 2U_0\sigma_1^2\right)U_T + \left(U_0^2\sigma_1^2 - U_1^2\sigma_0^2 + 2\sigma_0^2\sigma_1^2 \ln\frac{\sigma_0}{\sigma_1}\right) = 0. \tag{8.14}$$

Solving for U_T in Eq. (8.14), the optimal decision threshold [54] corresponding to the lowest bit-error rate is

$$U_{OT} = \frac{\sigma_1\sigma_0}{\sigma_1^2 - \sigma_0^2}\sqrt{(U_1 - U_0)^2 + 2\left(\sigma_1^2 - \sigma_0^2\right)\ln\left(\frac{\sigma_1}{\sigma_0}\right)} + \frac{U_0\sigma_1^2 - U_1\sigma_0^2}{\sigma_1^2 - \sigma_0^2}. \tag{8.15}$$

Under the influence of atmospheric turbulence, the optimal threshold changes with time. According to Eq. (8.15), the threshold can be changed from the perspective of signal and variance tracking, and the minimum bit-error rate can be obtained theoretically.

Typical Adaptive-Threshold Detection Models

LMS Filter Adaptive-Threshold Detection Model

The least-mean-square (LMS) algorithm [55] is based on the most jittery gradient-descent method, which includes two basic processes: filtering and adaptive. The filtering process calculates the response of the linear-filter output to the input signal. By comparing the output result y_k with the expected response d_k, the estimated error ε_k ($\varepsilon_k = d_k - y_k$) is generated, and the error signal ε_k is fed back to the filter. The adaptive process automatically adjusts the filter parameters according to the feedback error to make the mean-square-error signal close to zero to obtain the best prediction value of the signal mean and variance.

The LMS filtering adaptive-threshold detection model is applicable to stable input signals. It is illustrated in Fig. 8.4. The difference between the output signal and the expected signal is used as the error signal. The error signal is controlled by the adaptive weight control algorithm, and the transverse-filter weight vector is adjusted to make the mean-square value of the error signal infinitely close to zero. This obtains the best predicted value of the signal mean and variance. Then, the value is substituted into the maximum-likelihood ratio expression of the receiver, and the optimal decision threshold is obtained.

According to the LMS adaptive-threshold detection model, the output signal y_k and transfer function H_k of the system are expressed as [56]

$$y_k = W_k^T X_k \tag{8.16}$$

$$H_k = \sum_{i=1}^{l} w_{ik}\delta(t - iT), \tag{8.17}$$

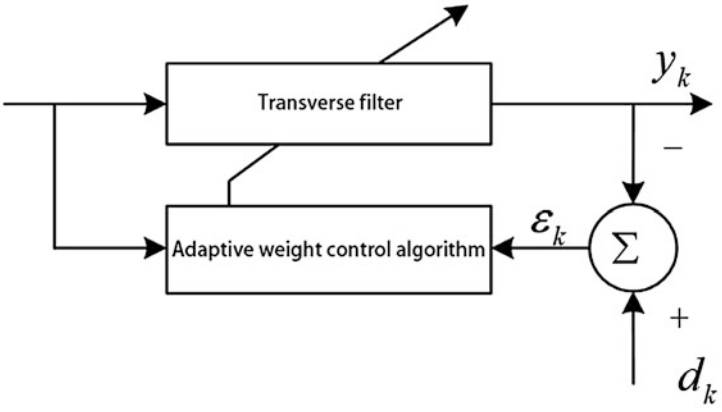


Fig. 8.4 LMS filter algorithm model [56]

where X_k is the input vector of the filter, W_k is the weight vector, l is the horizontal filter length, w_{ik} is the weight coefficient, and T is the symbol duration.

The output signal y_k is substituted into the error-signal ε_k expression, $\varepsilon_k = d_k - W_k^T X_k$, and its mean-square value is calculated to obtain the mean-square-error performance surface [56]:

$$\xi = E\left(d_k^2\right) + W_k^T R W_k - 2PW_k, \tag{8.18}$$

where R is the input signal X_k correlation matrix and P is the cross-correlation matrix between the expected signal d_k and the input signal y_k. The gradient of the mean-square error performance surface is calculated by $\nabla_k = \partial\xi_k/\partial W_k = 2RW_k - 2P$ because ε_k^2 is used as an estimate of ε_k in the LMS filtering algorithm:$\widehat{\nabla}_k = -2\varepsilon_k X_k$.

The LMS algorithm weight-vector update equation [56] is expressed as

$$W_{k+1} = W_k + 2\mu_c \varepsilon_k X_k \tag{8.19}$$

where W_k is the initial weight vector, μ_c is the step size of the LMS, which is used to control the algorithm stability and convergence rate, and ε_k is the error between the output signal and the expected signal.

Adaptive-threshold detection based on an LMS filter does not need to clarify the statistical relationship between the noise and signal, and the algorithm is easy to implement. However, the weight adjustment process of the LMS filter is prone to nonconvergence and a large amount of computation, which leads to an increase in the bit-error rate of the communication system.

Kalman-Filter Adaptive-Threshold Detection Model

The Kalman filter [55] is an optimal estimation algorithm that uses the linear-system state equation to combine the past measurement-estimation error with the new measurement error to estimate the future error. It is generally used to process nonstationary input signals. The working principle is illustrated in Fig. 8.5. The system state x_k at time k is obtained by multiplying the system-disturbance noise vector w_k with the system state x_{k-1} at time $k-1$ and the state-transition matrix A_{k-1}.

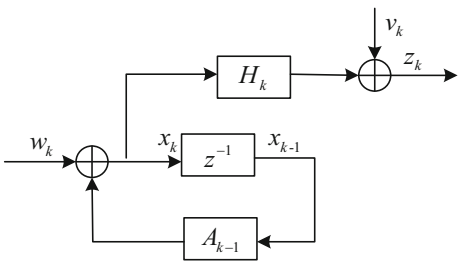


Fig. 8.5 Kalman-filtered signal model [57]

The observation value z_k at time k can be obtained by multiplying x_k with the observation matrix H_k and adding the observation noise vector v_k.

An adaptive threshold based on the Kalman filter is often used to solve prediction problems. It estimates the current value of the state variable using the estimated value of the previous state and the current observation value. It then compares the current value with the decision threshold calculated by the previous symbol, gives the decision result, and updates the detection threshold. Essentially, it is a group of optimal predictor–corrector estimators. The best approach is to minimize the covariance of the estimated error when certain limited conditions are met.

The system-state and measurement equations are established according to the Kalman adaptive-threshold detection signal model [55]:

$$x_k = A_{k-1}x_{k-1} + w_k \tag{8.20}$$

$$z_k = H_k x_k + v_k. \tag{8.21}$$

The gain K_k of the Kalman filter is defined as [58]

$$K_k = P_{k|k-1}{H_k}^T \left(H_k P_{k|k-1}{H_k}^T + R_k\right)^{-1}, \tag{8.22}$$

where R_k is the covariance matrix of the measurement noise vector v_k, $P_{k|k-1} = A_{k|k-1}P_{k-1}A_{k|k-1}^T + Q_k$ is the covariance matrix of the one-step prediction error, and Q_k is the covariance matrix of the process noise vector w_k. From this, the predicted value of the signal [58] is

$$\widehat{x}_k = \widehat{x}_{k|k-1} + K_k\left(z_k - H_k\widehat{x}_{k|k-1}\right). \tag{8.23}$$

The prediction of the fluctuation variance of the logarithmic amplitude of the turbulence [58] can be expressed as

$$\widehat{\sigma}^2{}_\chi = \frac{1}{4}\left(\widehat{\sigma}^2{}_k + 1\right), \tag{8.24}$$

where $\widehat{\sigma}^2{}_k$ is the estimated value of the signal variance at time k. According to the maximum-likelihood ratio expression, the decision threshold of the receiver is related to the amplitude fluctuation variance of the number. By knowing the estimated value of the amplitude-fluctuation variance, the maximum-likelihood ratio can be obtained for adaptive-threshold detection.

Adaptive-threshold detection based on the Kalman filter avoids the online calculation of the system transfer function; the transfer function can be calculated offline. This simplifies the calculation of the filter and reduces the computational burden, making it easier to implement in hardware, and the real-time performance of the filter is better. However, the relationship between noise and the

signal needs to be clarified, which is applicable to nonstationary random processes.

Adaptive-Threshold Detection Model with a Fading Kalman Filter

As shown in Fig. 8.5, the traditional Kalman filter estimates the current state based on an estimated value of the previous state and the current observation value. When the observation model is inaccurate, the old observation value has a significant impact on the current-state estimation and the role of the current observation data is ignored, which causes a filtering error. Adaptive-threshold detection with a fading-memory Kalman filter can solve the above problems.

This method adds a fading factor λ_k to the traditional Kalman filter so that the prediction-variance matrix [57] is

$$P_{k|k\text{-}1} = \lambda_k \left(A_{k|k\text{-}1} P_{k\text{-}1} A_{k|k\text{-}1}^T + Q_k \right). \tag{8.25}$$

In the formula, λ_k is the fading factor; it will reduce the state-filtering performance if it is too large. $A_{k|k-1}$ is the one-step state-transition matrix, P_{k-1} is the k_1-moment error covariance matrix, and Q_k is the correlation matrix of the process noise vector.

The estimated values of the prediction-error and innovation covariances [33] are as follows:

$$\widehat{P}_{k|k\text{-}1} = \lambda_k P_{k|k\text{-}1} \tag{8.26}$$

$$\widehat{C}_k = \alpha_k C_k, \tag{8.27}$$

where α_k is a scalar factor, $\widehat{C}_k$ is determined by the windowing estimation method, and the fading factor λ_k is solved. Then, the signal state filtering is [33]

$$\widehat{x}_k = A_{k|k-1} \widehat{x}_{k|k-1} + K_k \left(z_k - H_k A_{k|k-1} \widehat{x}_{k|k-1} \right). \tag{8.28}$$

The estimated logarithmic fluctuation variance is the same as that of adaptive-threshold detection with a traditional Kalman filter. The predicted value of the logarithmic amplitude-fluctuation variance is calculated using the predicted signal-variance value at time k, and the likelihood ratio of the receiver is obtained for threshold judgment. The adaptive threshold based on a fading Kalman filter reduces the filter memory length and causes the current data to play a greater role in the filtering process.

These three adaptive-threshold detection models are applicable to adaptive signal processing under weak-turbulence conditions. For strong-turbulence conditions, adaptive-threshold detection technology based on high-order cumulants is typically used.

High-Order Cumulant Adaptive-Threshold Detection

In recent years, Professor Ke Xizheng of Xi'an University of Technology has extensively researched adaptive optical systems in optical wireless communications, including adaptive algorithms, blind equalization, and threshold detection technology with high-order cumulants [24, 25, 31, 61, 62].

The noise model in optical wireless communication systems typically follows a Gaussian distribution and generally uses second-order statistics (correlation function and power spectrum) to analyze and process the noise. For non-Gaussian signal processing, higher-order (third and higher) cumulants can be used. Adaptive-threshold detection based on higher-order cumulants can estimate channel parameters and noise, better suppress the impact of strong turbulence on the received signal, improve the system communication quality, and is suitable for correlated, partially correlated, and uncorrelated signal models.

Decision-Guided Adaptive-Filtering Algorithm

The decision-guided algorithm has a weak convergence, but its steady-state tracking performance is good. Therefore, the decision-guided algorithm was combined with the LMS adaptive-filtering algorithm based on higher-order cumulants [24]. The decision-guided adaptive-filtering algorithm based on higher-order cumulants uses higher-order cumulants as the cost function to update the tap coefficients of the filter. This not only obtains a small mean-squared error but also effectively suppresses additive noise and symmetrically distributed non-Gaussian colored noise.

Figure 8.6 shows the mean-squared error curve when the higher-order cumulant decision-guided algorithm and the decision-guided algorithm select the optimal step size. As shown in Fig. 8.6, compared with the decision-guided algorithm, the higher-order cumulant decision-guided algorithm converges slowly, but its steady-state error is improved.

In addition, the error performance of the high-order cumulant decision-guided adaptive-filtering algorithm was studied for different noise types in the atmospheric channel, and the mean-squared error curves of different step sizes and those of the three noise models under the same step size were obtained, as shown in Figs. 8.7 and 8.8, respectively. As shown in Fig. 8.7, the larger the step size, the faster the curve convergence speed, but the larger the steady-state error.

As shown in Fig. 8.8, under the additive-noise condition, the convergence speed of the algorithm is fast and the steady-state error is small. With multiplicative noise, the convergence speed is obviously slowed down, and the steady-state error is also relatively increased. Multiplicative noise converges faster than combined multiplicative and additive noise, but the steady-state error difference between them is small.

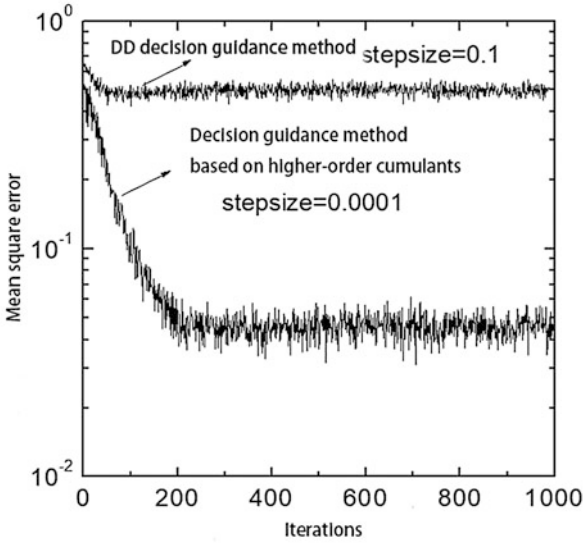


Fig. 8.6 Performance comparison of two decision-guided algorithms [24]

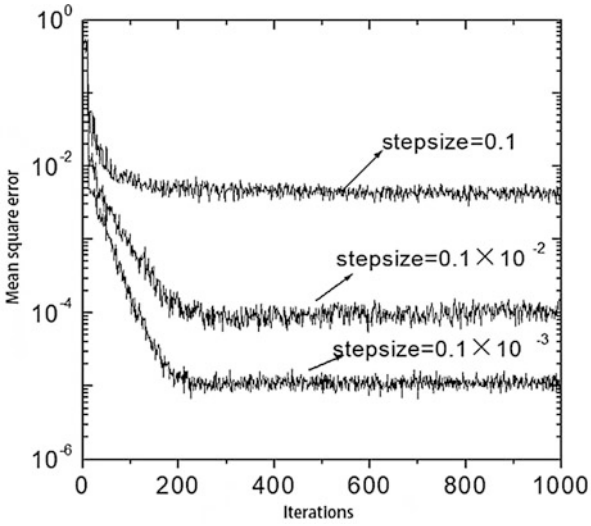


Fig. 8.7 Mean-square error (MSE) curve of asynchronous lengths [24]

Sliding-Correlation Adaptive-Threshold Algorithm

Sliding-correlation detection [59] refers to the detection or denoising of signals through autocorrelation and cross-correlation operations. Generally, the sliding-window method is used to judge the presence or absence of each window signal.

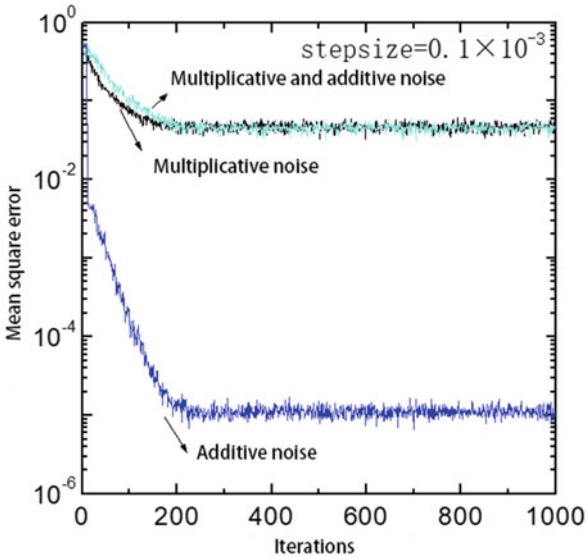


Fig. 8.8 Mean-square error (MSE) curve of three noise models [24]

Because the correlation function only considers the observed additive-white-noise data, it is proposed to combine the higher-order cumulant with correlation detection.

The detection principle is illustrated in Fig. 8.9. When a local pulse signal is input, the phase of the signal carrier is random; therefore, it needs to be nonlinearly transformed ($[\cdot]^4$ represents a fourth-power nonlinear transformation). The fourth harmonic component is extracted, and then the signal is detected using a digital filter and a sliding-correlation algorithm. This method can achieve multiple correlations and improve the detection probability of the system (Fig. 8.9).

Sliding-correlation adaptive-threshold detection based on higher-order cumulants [25] performs a cumulative correlation operation on the received signal according to the sliding-correlation principle. There are L correlation peaks in each frame of length L of the output signal, of which only one is the maximum autocorrelated peak carrying the transmitted signal information. The adaptive-threshold control algorithm then judges, accumulates, and compares the cumulative-output correlation peaks within a given observation frame length and obtains the adaptive decision threshold. The adaptive-threshold control model is shown in Figs. 8.10 and 8.11.

The correlation-peak sample value U_j of the threshold decision box is input, and the sample value of its adjacent signal frames U_{j-M+1}, U_{j-M}, …, U_j is used as the observation object (M is an integer). The threshold decision maker determines the maximum autocorrelation peak obtained by the correlation operation in each frame period. If $U_{j-M} \geq \lambda_j$, the decision maker outputs $Z(j) = 1$; if $U_{j-M} < \lambda_j$, the decider outputs $Z(j) = 0$. The accumulator accumulates the sample values of the signal output $N \times L$ $Z(j)$ by the decision, and the output signal is $X_j = \sum_{M=0}^{M-1} Z(j-M)$.

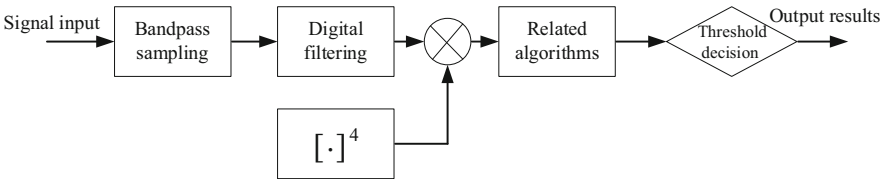


Fig. 8.9 High-order cumulant sliding-correlation detection [25]

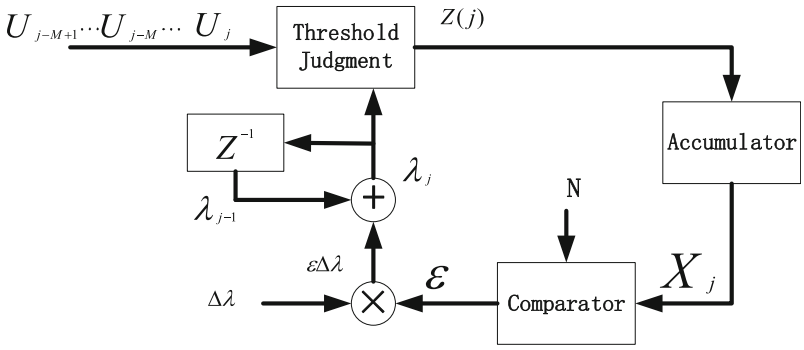


Fig. 8.10 Adaptive-threshold control algorithm for sliding correlation detection [25]

In the comparator, the sum of accumulation N and X_j is compared with the predetermined number of reference symbols to obtain a value for variable ε to control the threshold-adjustment direction. The comparison process is equivalent to a symbol function, that is, $\varepsilon = SGN(X_j - N)$. The threshold change amount $\Delta\lambda$ is obtained by multiplying it by the threshold control increment $\varepsilon\Delta\lambda$. Finally, adaptive control with the decision threshold $\lambda_j = \lambda_{j_1} + \varepsilon\Delta\lambda$ is realized.

Adaptive-Threshold Detection in a Strong-Turbulence Channel

Optical wireless communication divides atmospheric channels into weak-, medium-, and strong-turbulence channels, according to the degree of atmospheric turbulence and state of motion [60]. A strong-turbulence channel refers to an atmospheric channel where the atmospheric refractive-index structure constant $C_n^2 > 2.5 \times 10^{-13} m^{-2/3}$.

For the adaptive-threshold detection of strong-turbulence channels, the Xi'an University of Technology has established a K-distribution strong-turbulence channel model [30] using a modified spherically invariant random-process method to avoid the ergodic problem of random variables. Based on this channel model, the high-order cumulant of the innovation process obtained from the light-intensity signal is analyzed using the protection characteristics between the high-order and low-order

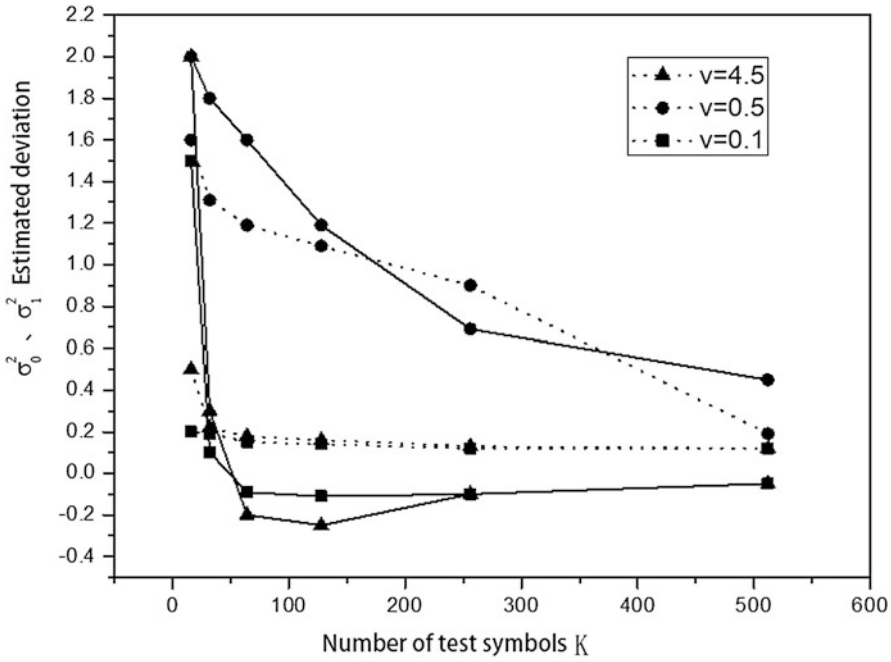


Fig. 8.11 Relationship between the estimated deviation and number of test symbols [31]

cumulants [31, 61]. Thus, the signal fading and Gaussian noise are estimated and then separately processed to obtain the real-time parameters of the turbulence signal to control the signal decision threshold.

Under OOK modulation, the signal-to-noise ratio $SNR = \mu/2\sigma_n^2$ of the system is defined under this K distribution and the signal-fading noise ratio $\mathrm{CNR_K}$ and two types of noise ratio $\mathrm{CNR_G}$ are set to 10 dB. The channel and noise correlation coefficients are set to 0.9, and the sample number $m = 32$. The deviation and variance between the noise parameters and theoretical values are studied [31].

Figures 8.11 and 8.12 show the estimated deviation and variance-change curves with theoretical values when using higher-order cumulants to estimate Gaussian noise parameters under different order-parameter v conditions, $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$. $\sigma_{n_1}^2$ is represented by a dotted line, and $\sigma_{n_2}^2$ is represented by a solid line. As shown in Figs. 8.11 and 8.12, when $v = 0.1$, the deviation and variance of the estimated value are the largest. With an increase in the order parameter v, the deviation and variance of the estimated value become increasingly smaller, indicating that the system tends to stabilize when the order parameter increases. However, the deviation and variance of the estimated value will not always decrease with an increase in v but will reach a critical point and tend to saturate.

When sampling at 50% of the pulse of the received digital signal, the mean value $\mu = 1$, order parameter $v = 2$, and the fixed threshold are set to 0.5. The system BER performance is studied when different detection methods are used, as shown in

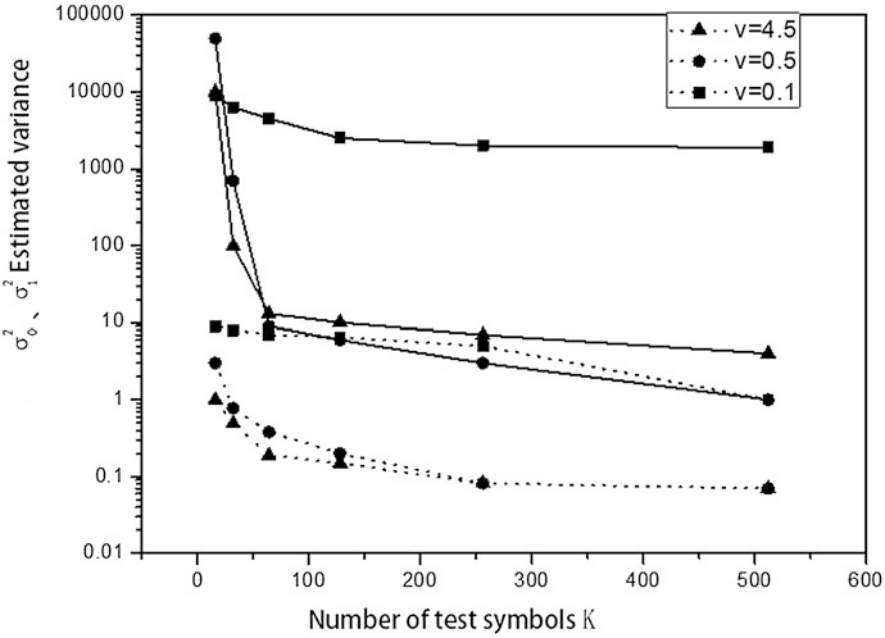


Fig. 8.12 Relationship between the estimated variance and number of test symbols [31]

Fig. 8.14. The dotted line represents the fixed-threshold detection method, and the solid line represents the adaptive-threshold detection method, when the sampling numbers *m* are 8, 16, 32, and 64.

As shown in Fig. 8.13, the bit-error rate of adaptive-threshold detection is significantly lower than that of the fixed-threshold detection method at the same signal-to-noise ratio. Moreover, the larger the number of samples in the adaptive-threshold detection, the more obviously the system performance improves.

Figures 8.14, 8.15, and 8.16 show the BER change curves with the SNR under heavy, moderate, and light rain, respectively. It can be seen that under the same SNR, the larger the number of samples, the lower the system BER and the more obvious the performance improvement of the adaptive-threshold detection method. When the number of samples exceeds 32, the performance improvement tends to flatten.

Figure 8.17 shows the change curves of the received bit-error rate with the signal-to-noise ratio in heavy, moderate, and light rain when the number of samples is 64. Under the same signal-to-noise ratio, the bit-error rate of the received signal in heavy rain is lower than that in moderate and light rain. They significantly improve the quality of the received signal and decrease the bit-error rate.

Based on the above adaptive-threshold detection experiments on rainy days, the error-rate performance of adaptive-threshold detection on sunny and foggy days was further studied [62]. The BER change curves with SNR on sunny, foggy, and rainy days are shown in Figs. 8.18, 8.19, and 8.20, respectively. The same conclusions were obtained from Figs. 8.18, 8.19, and 8.20 as those in

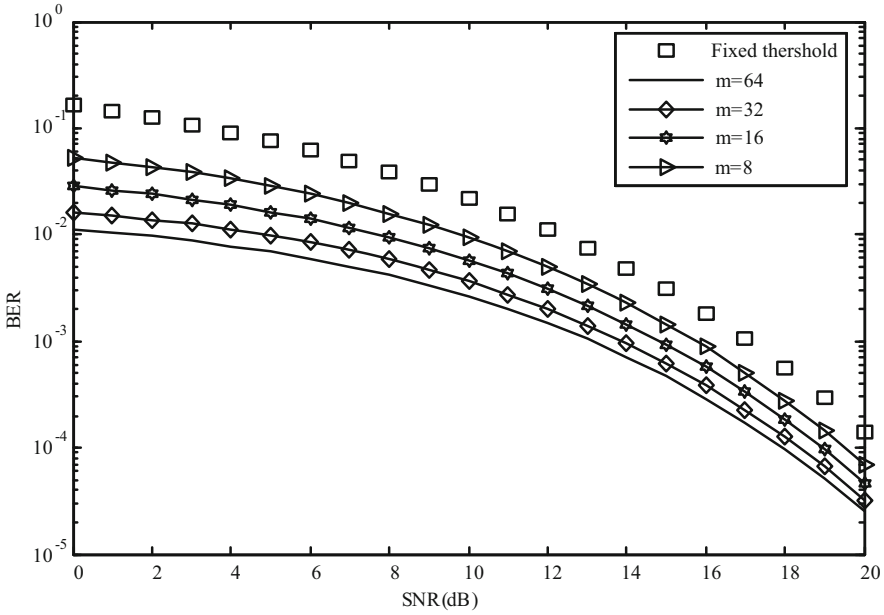


Fig. 8.13 BER change curve with SNR [61]

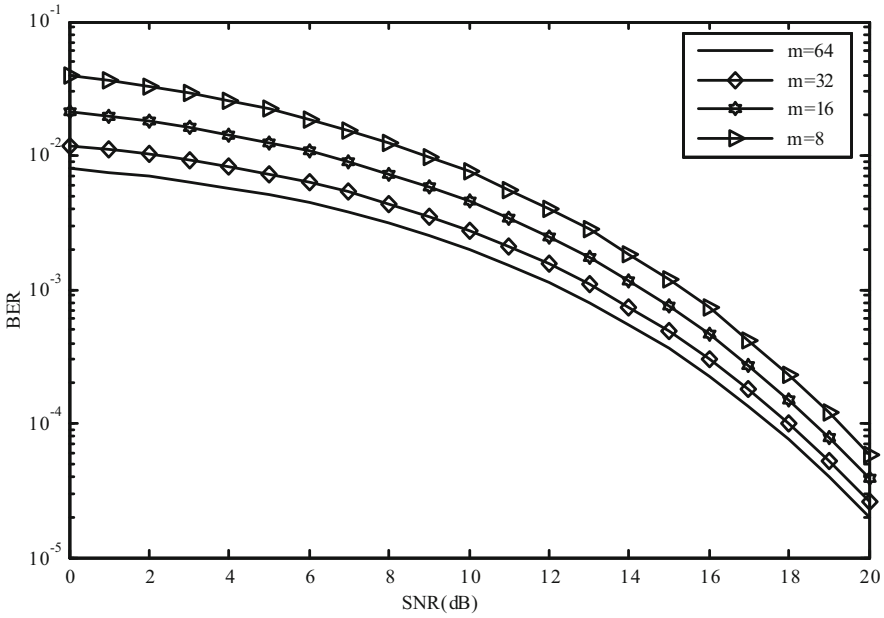


Fig. 8.14 Change curve of the bit-error rate with SNR in heavy rain [31]

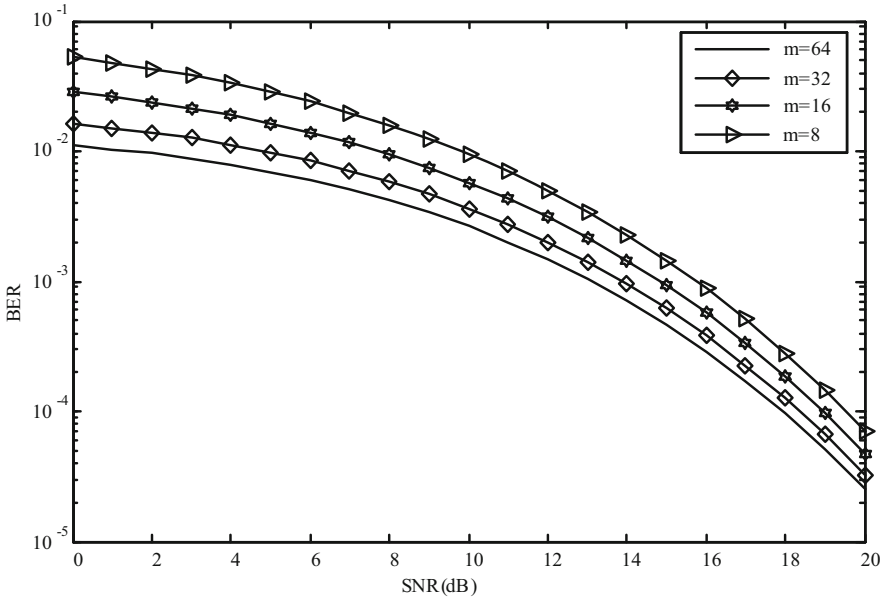


Fig. 8.15 Change curve of the bit-error rate with SNR in moderate rain [31]

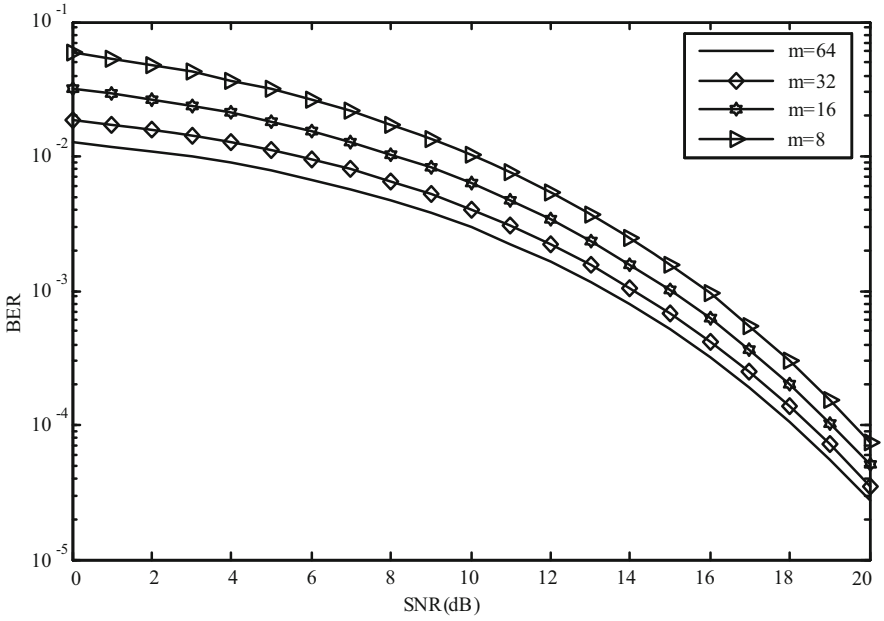


Fig. 8.16 Change curve of the bit-error rate with SNR in light rain [31]

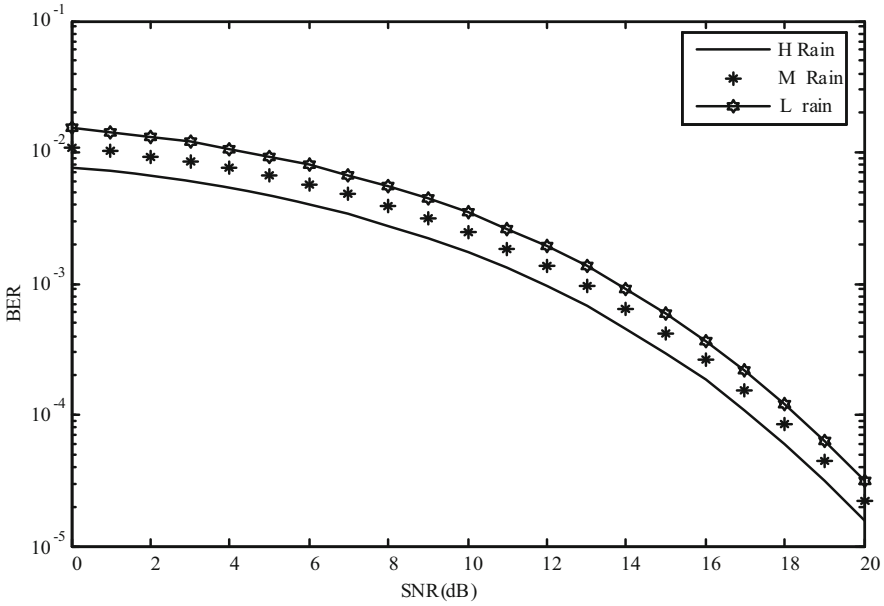
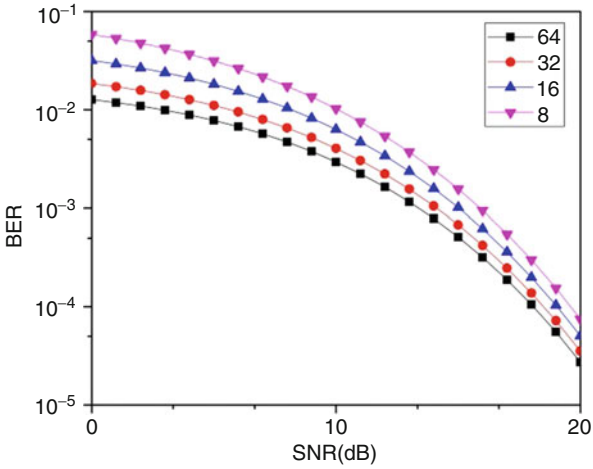


Fig. 8.17 BER change curve with SNR in different weather conditions [31]

Fig. 8.18 Relationship between the SNR and BER on sunny days [62]



Figs. 8.14, 8.15, and 8.16, which again verifies the correctness of the relationship between the sample number and BER.

Figure 8.21 shows the change curves of the received-signal error rate with the signal-to-noise ratio under different weather conditions. Under the same sampling value, the signal-to-noise ratio is inversely proportional to the error rate, indicating that increasing the signal-to-noise ratio can improve the communication quality.

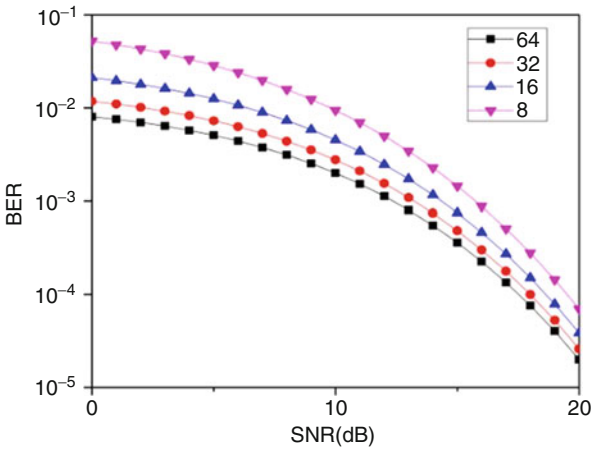


Fig. 8.19 Relationship between SNR and BER in foggy weather [62]

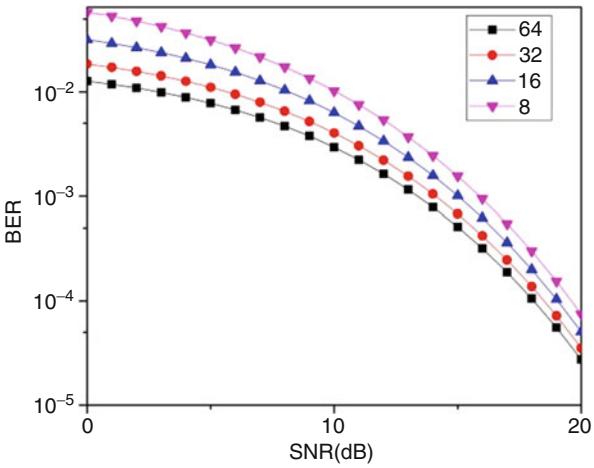


Fig. 8.20 Relationship between SNR and BER on rainy days [62]

Summary

The performance of an atmospheric optical wireless communication system is significantly affected by its decision threshold. Compared with traditional fixed-threshold detection, adaptive-threshold detection can update the optimal decision threshold in real time, improve the detection probability, shorten the acquisition time, and effectively reduce the bit-error rate at a lower cost.

This chapter reviewed the research progress of adaptive-threshold detection and introduced the optimal threshold strategy and several commonly used adaptive-threshold detection algorithms. Moreover, it discussed the development trends of

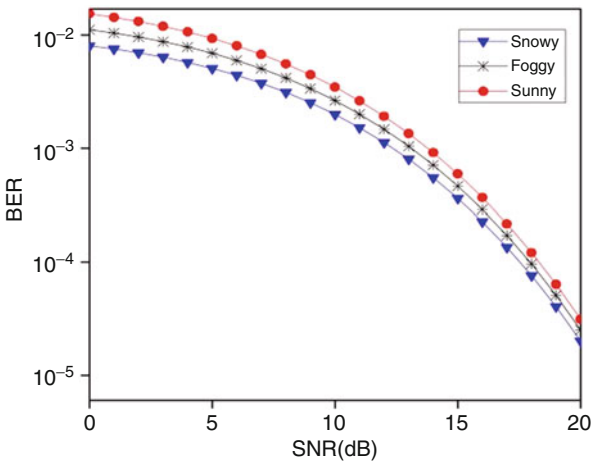


Fig. 8.21 Relationship between SNR and BER under different weather conditions [62]

adaptive-threshold detection technology in optical wireless communication, including the following aspects:

(1) As the data-transmission rate of optical wireless communication progressively increases, it is difficult to obtain channel-state information in high-speed optical communication. Therefore, it is necessary to study threshold detection techniques without channel-state information, such as blind estimation.
(2) Adaptive-threshold control algorithms (such as maximum likelihood, generalized maximum likelihood, and minimum mean square error) should be optimized, the number of iterations and algorithm complexity should be reduced, and real-time threshold decisions of the received signal should be made.
(3) The optimal decision threshold was studied under different modulation formats and turbulence intensities to observe signal detection under complex atmospheric-channel conditions.
(4) Based on the relatively slow-varying characteristics of statistics such as average signal intensity and turbulence intensity, the optimal constant threshold detection technique is studied to reduce the difficulty of implementing adaptive threshold detection.

References

1. Khalighi M A, Uysal M. Survey on free space optical communication: A communication theory perspective [J]. IEEE communications surveys & tutorials, 2014, 16(4): 2231–2259.
2. Chen Mu, Ke Xizheng. Study on the influence of atmospheric turbulence on the performance of laser communication system [J]. Infrared and Laser Engineering,2016,45(08):115–121.
3. Choi K, Cheun K, Jung T. Adaptive PN code acquisition using instantaneous power-scaled detection threshold under Rayleigh fading and pulsed Gaussian noise jamming [J]. IEEE Transactions on Communications, 2002, 50(8): 1232–1235.

4. Burris H R. Adaptive signal processing methods for bit error rate reduction in optical communication systems [D]. Washington: The Catholic University of America, 2003.
5. Urkowitz H. Energy detection of unknown deterministic signals [J] IEEE, 1967, 55(4): 523–531.
6. Liu Naian, Zeng Xingwen. Adaptive threshold technology and its performance in high-speed burst communication [J]. Electronic journal,1998,26(1):111–114.
7. Mukai R, Arabshahi P, Yan T Y. An adaptive threshold detector and channel parameter estimator for deep space optical communications [J]. IEEE, 2001,1(3): 50–54.
8. Burris H R, Reed A E, Namazi N M, et al. Adaptive Thresholding for Free-space Optical Communication Receivers with Muliplicative Noise [J]. IEEE, 2002,3(3):1473–1480.
9. Zhu X M, Kahn J M, Wang J. Mitigation of turbulence-induced scintillation noise in free-space optical links using temporal-domain detection techniques [J]. IEEE Photonics Technology Letters, 2003, 15(4): 623–625.
10. Cole M, Kiasaleh K. Signal intensity estimators for free-space optical communications through turbulent atmosphere [J]. IEEE Photonics Technology Letters, 2004, 16(10): 2395–2398.
11. Namazi N M, Burris R, Gilbreath G C. Analytical approach to the calculation of probability of bit error and optimum thresholds in free-space optical communication [J]. Optical Engineering, 2007, 46(2):25–32.
12. Riediger M L B, Schober R, Lampe L. Decision-feedback detection for free-space optical communications [J]. IEEE, 2007,3(7): 1193–1198.
13. Riediger M L B, Schober R, Lampe L. Fast multiple-symbol detection for free-space optical communications [J]. IEEE Transactions on Communications, 2009, 57(4): 1119–1129.
14. Moradi H, Refai H H, LoPresti P G. Thresholding-based optimal detection of wireless optical signals [J]. Journal of Optical Communications and Networking, 2010, 2(9): 689–700.
15. Chatzidiamantis N D, Karagiannidis G K, Uysal M. Generalized maximum-likelihood sequence detection for photon-counting free space optical systems [J]. IEEE transactions on communications, 2010, 58(12): 3381–3385.
16. Song T, Kam P Y. A robust GLRT receiver with implicit channel estimation and automatic threshold adjustment for the free space optical channel with IM/DD [J]. Journal of lightwave technology, 2013, 32(3): 369–383.
17. Beaulieu N C, Bartoli G, Marabissi D, et al. The structure and performance of an optimal continuous-time detector for Laplace noise [J]. IEEE communications letters, 2013, 17(6): 1065–1069.
18. Dabiri M T, Sadough S M S. Generalized blind detection of OOK modulation for free-space optical communication [J]. IEEE Communications Letters, 2017, 21(10): 2170–2173.
19. Choi I Y, Shin W H, Han S K. CSI estimation with pilot tone for scintillation effects mitigation on satellite optical communication [J]. Optics Communications, 2019, 435(11): 88–92.
20. Shin W H, Choi J Y, Han S K. Fixed threshold on-off keying differential detection for satellite optical communications [J]. Optics Express, 2019, 27(2): 1590–1596.
21. Choi J Y, Lee J W, Han S K. Stochastic Approximation Aided Adaptive Thresholding for Optical Detection in PAM4 Based FSO Transmission [J]. IEEE Access, 2021, 9(7): 106451–106459.
22. Wang Jin, Huang Dexiu. Turbulence noise suppression based on LSM-Kalman filter [J]. Photoelectric engineering, 2006,33(7):39–43.
23. Wu Yucheng, Chen Tingting. Adaptive threshold detection method for decision in sliding window [J]. Journal of System Simulation, 2008,20 (10): 2770–2773.
24. Wang Ni. Research on blind equalization of atmospheric laser communication channel based on higher-order cumulants [D]. Xi'an: Xi'an University of Technology, 2009.
25. Zhao Zhifeng. Design and implementation of power line carrier centralized meter reading system [D]. Xi'an: Xi'an University of Technology, 20010.
26. Wang Z X, Zhong W D, Yu C. Performance improvement of OOK free-space optical communication systems by coherent detection and dynamic decision threshold in atmospheric turbulence conditions [J]. IEEE Photonics Technology Letters, 2012, 24(22): 2035–2038.

27. Li Fengfei, Hu Guorong. Adaptive timing synchronization threshold selection method for power line OFDM communication system [J]. Power System Automation, 2013,37 (23): 93–96.
28. Shang Yaobo, Guo Ying, Wang Jinjiang, et al. Adaptive threshold decision PN code acquisition algorithm based on SNR [J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2014,34 (04): 88–93.
29. Zhang Bo, Guo Ying, Hou Wenlin, et al. Fast adaptive threshold detection algorithm for variable threshold factor [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2015,43 (08): 62–65.
30. Ke Xizheng, Tian Xiaochao. Modeling and simulation of two-dimensional correlated K-distribution turbulent channel [J]. Journal of Optics, 2015,35 (04): 41–410.
31. Tian Xiaochao. Adaptive processing of received signals in atmospheric laser communication under strong turbulence [D]. Xi'an: Xi'an University of Technology, 2018.
32. Ding S L, Zhang J K, Dang A H. Adaptive threshold decision for on-off keying transmission systems in atmospheric turbulence [J]. Optics express, 2017, 25(20): 24425–24436.
33. Zhang Hongling, Liu Zhi, Ni Xiaolong, et al. Research on optimization technology of adaptive threshold in atmospheric optical communication [J]. Information Communication, 2018, 185 (05): 229–230.
34. Li Xiaoyan, Zhang Peng, Tong Shoufeng. BER performance analysis of reverse modulation free-space optical communication system based on adaptive decision threshold under the influence of atmospheric turbulence [J]. China Laser, 2018,45 (06): 161–170.
35. Zhu Baoqin, Zhang Mingqing, Chen Jian. Optimization of OOK/FSO system performance by adaptive decision threshold and DC extraction technology [J]. Journal of Quantum Electronics, 2018,35 (01): 108–116.
36. Li L, Geng T W, Wang Y, et al. Free-space optical communication using coherent detection and double adaptive detection thresholds [J]. IEEE photonics journal, 2018, 11(1): 1–18.
37. Li X Y, Zhao X H, Zhang P. Bit error rate analysis for modulating retro-reflector free space optical communications with adaptive threshold over correlated gamma gamma fading channels [J]. IEEE Communications Letters, 2019, 23(12): 2275–2279.
38. Jiang X L, Diao M. A New Type Double-Threshold Signal Detection Algorithm for Satellite Communication Systems Based on Stochastic Resonance Technology [J]. Wireless Networks,2019,12(13):259–272.
39. Zhang Shengkui, Yao Zhicheng, Fan Zhiliang, et al. Frequency hopping signal extraction and detection based on local adaptive threshold [J]. Electro-optics and Control, 2020,27 (01): 68–72.
40. Liu Yang. An adaptive threshold carrier lock detection algorithm for 16QAM [J]. Telecommunication Technology, 2020,60 (02): 164–169.
41. Liu C, Xie F, Dong X, et al. Small Target Detection From Infrared Remote Sensing Images Using Local Adaptive Thresholding [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15(1): 1941–1952.
42. Wang Huiqin, Wang Xue, Sun Jianfeng, et al. Average capacity of wireless optical multiple-input multiple-output system under correlated channels [J]. Journal of Optics, 2016,36 (07): 41–410.
43. Ke Xizheng, Deng Lijun. Optical Wireless Communication [M]. Beijing: Science Press, 2016.
44. Lu Hui, Gao Zhan. Laser and detector in free space optical communication system [J]. Telecom Express, 2003,15 (03): 35–38.
45. Li Fei, Lu Houbing. Threshold optimization method for atmospheric optical communication under weak turbulence conditions [J]. Infrared and Laser Engineering, 2016,45 (12): 190–195.
46. Wu Changluan, Xiang Jianyong, Ma Jingyi, et al. Photoelectric accumulation detection principle and test method of detection probability and false alarm probability [J]. Photoelectric Engineering, 1990,17 (05): 26–34
47. Middleton D. Canonically optimum threshold detection [J]. IEEE Transactions on information Theory, 1966, 12(2): 230–243.

48. Arnon S. The effects of atmospheric turbulence and building sway on optical wireless communication systems [J]. Optics Letters, 2003, 28(2):129–131.
49. Wang Jin. Research on adaptive reception technology for optical wireless communication [D]. Wuhan: Huazhong University of Science and Technology, 2006.
50. Lin Jing. Wavelet threshold denoising technology based on maximum likelihood estimation and signal feature extraction [J] Journal of Instrumentation, 2005,26 (09): 923–928.
51. Zhang Xianda. Modern Signal Processing [M]. Beijing: Tsinghua University Press, 2002.
52. Lee P. Computation of the bit error rate of coherent M-ary PSK with Gray code bit mapping [J]. IEEE Transactions on Communications, 1986, 34(5): 488–491.
53. Wang J, Yuan X, Huang D. Suppression of turbulence noise in optical wireless communication system [J]. SPIE, 2005, 5625(11): 1077–1084.
54. Gagliardi R M, Karp S. Optical Communications [M]. New York: Wiley-Interscience, 1995.
55. Simon Haykin. Adaptive Filter Theory [M]. Upper Saddle River:Prentice Hall, 2002.
56. Wang Jin, Huang Dexiu, Yuan Xiuhua. Analysis of reception performance of Optical Wireless Communication based on minimum mean square adaptive filter [J]. China Laser, 2006,33 (10): 1379–1383.
57. Zhang Hongling. Research on atmospheric optical communication optimization technology based on adaptive threshold [D]. Changchun: Changchun University of Technology, 2018
58. Wang Jin, Huang Dexiu, Yuan Xiuhua, et al. Atmospheric turbulence noise suppression based on adaptive filter [J]. Optical Technology, 2006,32 (05): 750–753.
59. Li Rui, He Fuyun, Xia Yubao. Relevant detection principle and its application [J]. Journal of Hefei University of Technology, 2008,31 (4): 573–576.
60. Wu Xiaojun, Wang Hongxing, Li Bifeng, etc. Statistical characteristics analysis of atmospheric refractive index structure constant near the sea surface [J] Journal of Optics, 2015, (4): 11.
61. Ke X Z, Ji X K. Research on Adaptive Threshold of Received Signal in Communication System [J]. Optics and Photonics Journal, 2021, 11(1):1–12.
62. Ji Xukuan. Research on signal threshold detection and processing of Optical Wireless Communication system [D]. Xi'an: Xi'an University of Technology, 2021.

Four-Quadrant Detector Light-Spot Detection Principle

9

Contents

Introduction 270
Light-Spot Model and Detection Principle 270
Light-Spot Model 270
QD-Based Spot-Position Detection Principle 271
Four-Quadrant Detector-Based Spot-Position Detection 273
Foreign Research Progress 273
Domestic Research Progress 278
QD Spot-Detection Influence Factors 282
Effect of Different Light-Spot Modes on the QD Output 282
Effect of Different Spot Radii on the QD Output 282
Effect of Different Dead-Zone Widths on the QD Output 282
Effect of Different Background Lights on the QD Output 283
Effect of Different Signal-to-Noise Ratios on the QD Output 283
Effect of Different Positioning Algorithms on the QD Output 284
Spot-Detection Algorithms 285
Addition-Subtraction Algorithm 285
Diagonal Algorithm 285
Difference-Ratio-Sum Algorithm 286
Logarithmic Algorithm 286
Function-Fitting Methods 287
QD Light-Spot Detection System 290
Beam-Alignment Detection System 290
Missing-Spot Detection Test 290
System Alignment Test 290
QD Composite Tracking and Communication System 291
Micro- and Nano-laser-communication Terminal 292
Summary 293
References 294

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_9

Abstract

Spot detection technology plays an important role in modern scientific research and application. Due to its advantages of high detection sensitivity, simple signal processing, and strong anti-interference ability, four-quadrant detectors have been widely used in spot detection technology and are the key components in acquisition, tracking, and pointing systems. First, the spot models commonly used in spot detection and the principle of four-quadrant detector spot detection are introduced. Then, the research results of four-quadrant detector spot detection at home and abroad, as well as the factors affecting the accuracy of four-quadrant detector spot detection and the commonly used spot detection algorithms, are analyzed. The four-quadrant detector spot detection system is introduced, including the beam alignment detection system, four-quadrant detector tracking communication composite system, and micronano laser communication system. Finally, the development prospects of spot detection technology are prospected.

Keywords

Spot detection · Four-quadrant detector · Spot model · Spot detection algorithm

Introduction

Acquisition, pointing, and tracking (APT) technology is a research focus in wireless laser communication systems, and spot-detection technology is key to APT technology. Commonly used spot detectors are position-sensitive detectors, charge-coupled devices (CCDs), and four-quadrant detectors (QDs). Compared with position-sensitive detectors and charge-coupled devices, QDs have a high detection resolution and short response time and are therefore widely used.

This chapter introduces several common spot models used in spot detection and then analyzes and summarizes the detection principle, domestic and foreign research progress, influencing factors, detection algorithms, and applications around the four-quadrant detector. Finally, it discusses potential applications of spot-detection technology and its development prospects, according to the development status.

Light-Spot Model and Detection Principle

Light-Spot Model

Each input spot model is different, as is the output signal value of the QD. The spot model is divided into uniform and Gaussian distributions by energy and into circular, elliptical, and toroidal spots by shape. The ideal spot is a uniformly distributed circular spot with an energy-density distribution function of [1]

$$I(x,y) = \begin{cases} I_0 & (x-x_0)^2+(y-y_0)^2 \le R^2 \\ 0 & \text{otherwise} \end{cases}. \tag{9.1}$$

I_0 is the peak light intensity, R is the spot radius, and the coordinates of the spot center are (x_0, y_0). The spot output from the laser can be regarded as a circular spot with a Gaussian distribution; its energy-density distribution function is [2, 3]

$$I(x,y) = I_0 \exp\left(\frac{-2\left((x-x_0)^2+(y-y_0)^2\right)}{R^2}\right). \tag{9.2}$$

To reduce the experimental error in experimental analyses, the received light spot is approximated as a uniformly distributed elliptical light spot with an energy-density distribution function of

$$I(x,y) = \begin{cases} I_0 & \dfrac{x^2}{w_a^2}+\dfrac{y^2}{w_b^2} \le 1 \\ 0 & \text{Otherwise} \end{cases}. \tag{9.3}$$

I_0 is the peak light intensity, and w_a and w_b are the elliptical light spot's long and short axes, respectively. When the spot drifts, the blocked spot is viewed as a Gaussian-distributed elliptical spot with an energy-density distribution function of [4, 5]

$$I = I_0 \frac{2}{\pi w_a w_b} \exp\left(-2\left[\frac{x^2}{w_a^2}+\frac{y^2}{w_b^2}\right]\right). \tag{9.4}$$

In practice, when using a Cassegrain telescope to receive the spot, the secondary mirror obscures the center of the spot, and a ring-shaped spot is presented. The energy-density distribution function of a uniformly distributed annular spot is [6]

$$I = k \iint\limits_{S_2-S_1} I_0 \frac{2}{\pi R^2} \exp\left(-2\left(\left((x-x_0)^2+(y-y_0)^2\right)/R^2\right)\right) ds, \tag{9.5}$$

where k is the conversion factor, R is the radius of the outer ring spot, S_1 is the area of the inner-ring circle, and S_2 is the area of the outer-ring circle.

QD-Based Spot-Position Detection Principle

QDs are photovoltaic-type semiconductor detector devices. A schematic diagram is shown in Fig. 9.1. It consists of four quadrants, each of which has a photodiode or avalanche photodiode mounted on the photosensitive surface. The spot is positioned at the center of the device by the output current of the four quadrants, when the spot is irradiated on the QD [7].

When the laser spot is incident on the QD, the four quadrants output different photocurrents I_A, I_B, I_C, and I_D, and the offset is calculated to determine the spot-

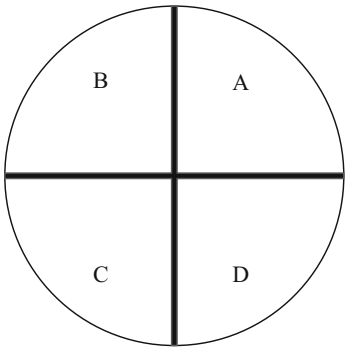


Fig. 9.1 Diagram of a QD

center position. The offsets on the σ_{x0} and σ_{y0} axes are represented by x_0 and y_0, and their equations using the conventional addition and subtraction algorithm are as follows:

$$\sigma_{x0} = (I_A + I_D) - (I_B + I_C) \tag{9.6}$$

$$\sigma_{y0} = (I_A + I_B) - (I_C + I_D). \tag{9.7}$$

Let the ratio coefficient between the QD output photocurrent and the corresponding quadrant spot energy be k. Then,

$$\sigma_{x0} = (I_A + I_D) - (I_B + I_C) = k[(E_A + E_D) - (E_B + E_C)] \tag{9.8}$$

$$\sigma_{y0} = (I_A + I_B) - (I_C + I_D) = k[(E_A + E_B) - (E_C + E_D)]. \tag{9.9}$$

E_A, E_B, E_C, and E_D denote the total energy of the spot incident on the four quadrants.

After normalizing σ_{x0} and σ_{y0}, we have

$$\sigma_x = \frac{(I_A + I_D) - (I_B + I_C)}{I_A + I_B + I_C + I_D} = \frac{(E_A + E_D) - (E_B + E_C)}{E_A + E_B + E_C + E_D} \tag{9.10}$$

$$\sigma_y = \frac{(I_A + I_B) - (I_C + I_D)}{I_A + I_B + I_C + I_D} = \frac{(E_A + E_B) - (E_C + E_D)}{E_A + E_B + E_C + E_D}. \tag{9.11}$$

Assuming a uniform spot-energy distribution, the spot area is proportional to the energy; then, we have

$$\sigma_x = \frac{(E_A + E_D) - (E_B + E_C)}{E_A + E_B + E_C + E_D} = \frac{(S_A + S_D) - (S_B + S_C)}{S_A + S_B + S_C + S_D} \tag{9.12}$$

$$\sigma_y = \frac{(E_A + E_B) - (E_C + E_D)}{E_A + E_B + E_C + E_D} = \frac{(S_A + S_B) - (S_C + S_D)}{S_A + S_B + S_C + S_D}. \tag{9.13}$$

S_A, S_B, S_C, and S_D denote the spot areas incident on the four quadrants.

Four-Quadrant Detector-Based Spot-Position Detection

Foreign Research Progress

Detecting the laser-spot position of a four-quadrant detector is a key technology in APT systems, and the spot-position detection accuracy directly affects the system performance. Many foreign scholars have conducted in-depth research on QD spot detection.

In 1985, the Japanese Space Agency developed the laser-utilizing communication equipment (LUCE) system [10]. The fine-tracking and overtargeting detectors of the APT subsystem of the LUCE system both used QDs, and the fine-tracking accuracy was better than 0.7 μrad.

Figure 9.2 shows the internal structure of the LUCE terminal [11], including the transmitting and receiving light paths. The receiving light path detects the spot position using the coarse-pointing mechanism sensor to drive the coarse-pointing mechanism for coarse alignment. The transmitting light path detects the change of the spot using the presighting sensor, which drives the presighting mechanism for alignment. After alignment, the transmitted beam is directed to the corner reflector, which reflects it to the fine-pointing sensor, which detects the change in the spot position.

In 1995, the Communication Research Laboratory (CRL) [12] of the Ministry of Posts and Telecommunications of Japan used Engineering Test Satellite VI to conduct a laser-communication test between the satellite-terminal laser communication equipment (LCE) and a ground station.

A block diagram of the LCE is shown in Fig. 9.3. The tracking module uses a QD with a 400-rad field of view, and the accuracy can be $\leq$2 μrad. The uplink is aligned by CCD and QD. The gimbal controller controls a two-axis gimbal mirror for coarse alignment, according to the spot information collected by the CCD. The fine-alignment controller controls the fine-aiming device for fine alignment according to the spot offset determined by the QD, and the aligned beam is irradiated to the communication Si avalanche photodiode (APD) through the polarizer.

The downlink is emitted by two lasers. After the beam expander, the beam is synthesized through the combined beam mirror into a single beam, and the beam direction is precisely controlled by a presighting device.

In 2001, Guelman [13] developed a broadband laser inter-satellite link (BLISL) system that uses QD to align and track the spot. The structure of the designed BLISL system is shown in Fig. 9.4. The communication beam is transmitted through the primary reflector, secondary reflector, fine-steering mirror, front-pointing mirror, and beam splitter to the receiver and quadrant detector. The fine-steering mirror is aligned according to the spot information detected by the quadrant detector.

The transmitting beam is transmitted through the transmitter, front-pointing mirror, and fine-steering mirror. The front-pointing mirror is used to compensate the beam deviation of the transmitter. The star tracker estimates the attitude of the gimbal with respect to the inertial reference system to minimize the aperture aiming error.

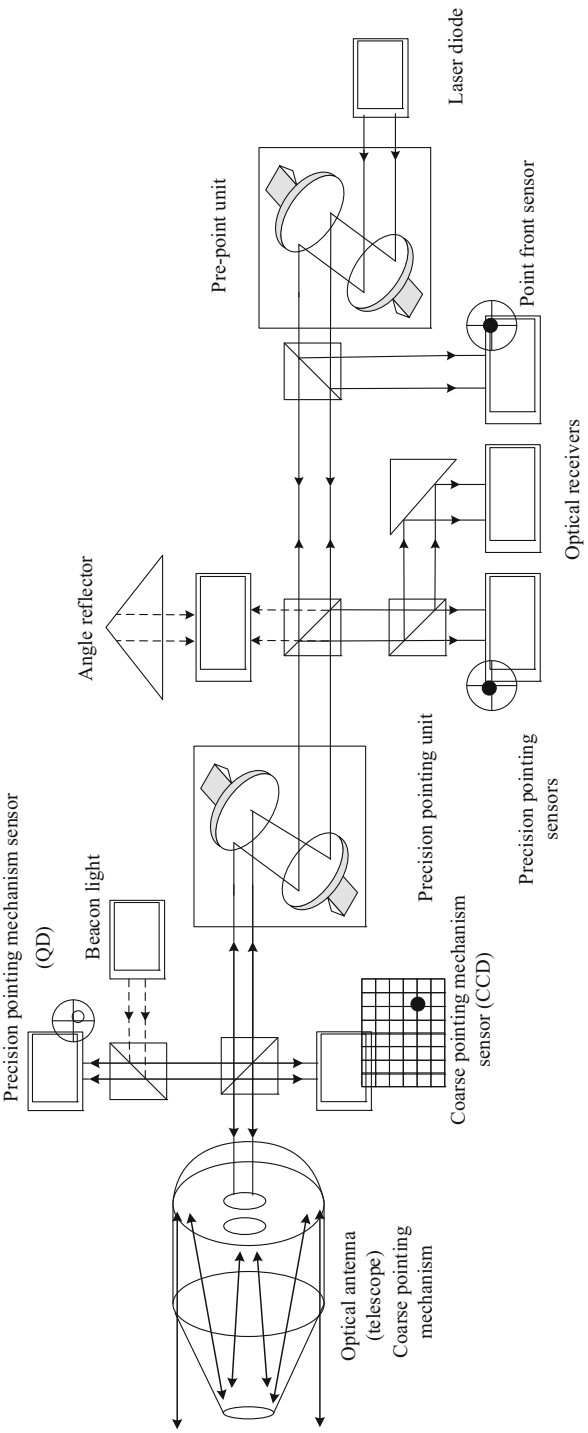


Fig. 9.2 Schematic diagram of the LUCE terminal [11]

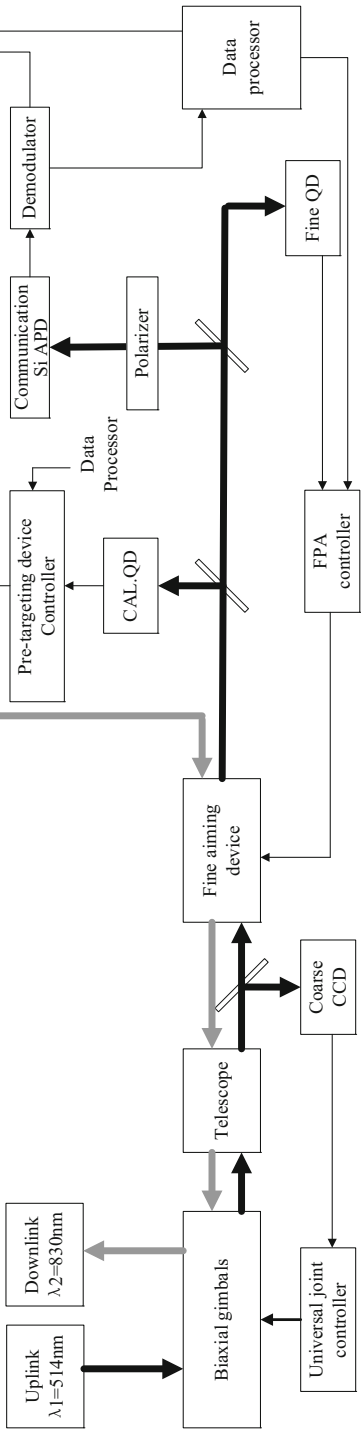


Fig. 9.3 LCE block diagram [12]

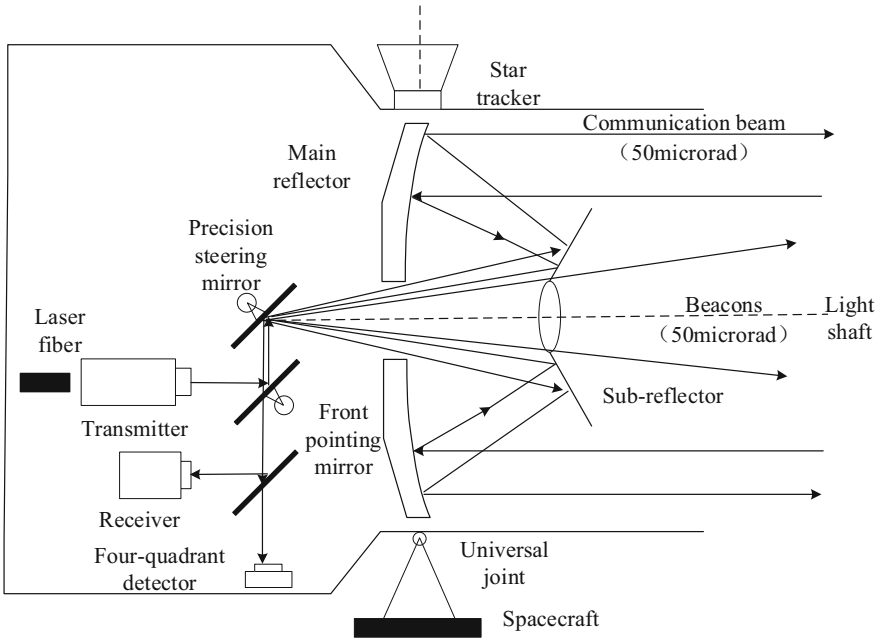


Fig. 9.4 BLISL system-architecture diagram [13]

In 2002, Toyoda et al. [8] performed a comparative study of the spot-detection sensor shown in Fig. 9.5 for a laser-tracking system. The received laser beam is focused on a quadrant APD, and a controller calculates the tracking signal and compensates the closed-loop system to drive the tilt mirror to keep the spot in the center of the four cells. An equivalent noise angle of less than 1 μrad is achieved at a received optical power greater than 4 pW.

In 2006, the Optical Inter-orbit Communication Engineering Test Satellite (OICETS) successfully conducted a two-way communication experiment with the optical ground-station terminal KODEN [14]. Figure 9.6 shows the KODEN receiver. The laser beam from OICETS is received and detected by a CCD camera, a QD sensor, and an avalanche photodetector (APD). QDs were used to detect and locate the spot position with a tracking accuracy of up to 2 μrad.

In 2009, Lee et al. [15] investigated the effect of the incident spot size on a four-quadrant photodetector. Experimental and analytical studies showed that the sensitivity of the optical displacement detection increased as the Gaussian spot size decreased.

The experimental setup is shown in Fig. 9.7. It consists of a beam expander, objective lens, mirror, and quadrant photodetector mounted on a linear-translation stage. After the beam expander, a 3-mW He–Ne laser with a 632-nm wavelength is focused by the objective mirror onto the reflector and reflected to reach the quadrant photodetector. The spot size is varied by changing the distance between the two lenses, L_1 and L_2, of the beam expander.

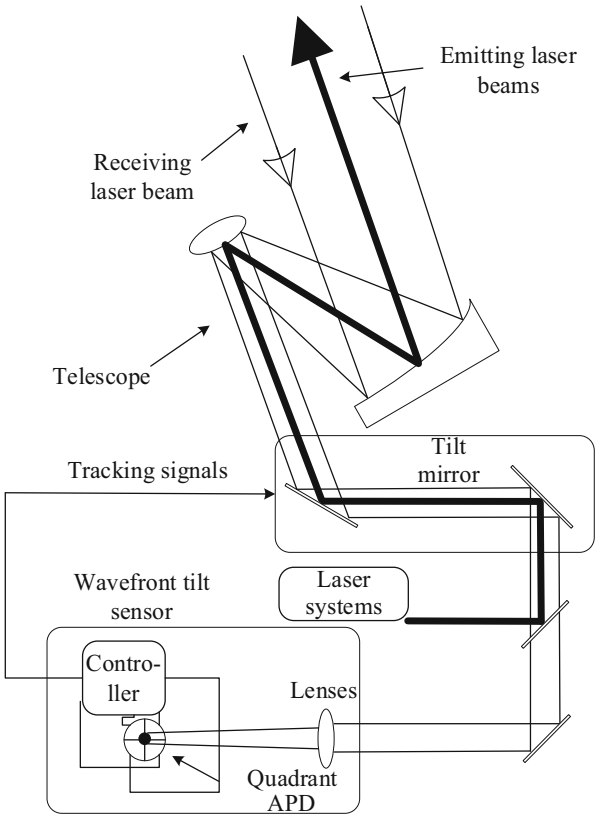


Fig. 9.5 Quadrant APD-based laser-tracking system [8]

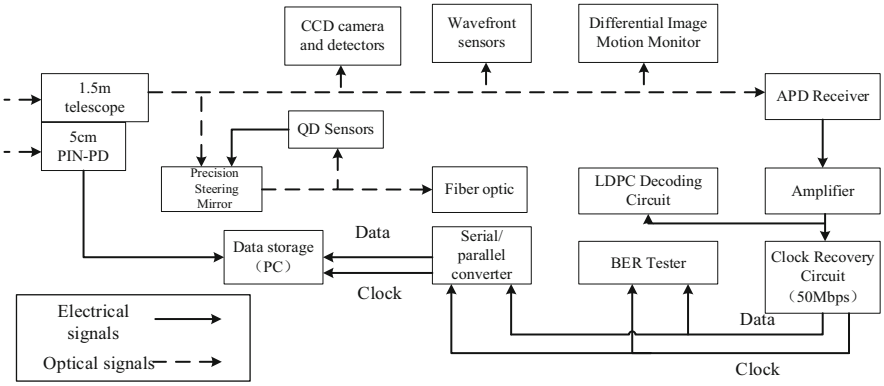


Fig. 9.6 KODEN receiver [14]

In 2012, Horwath et al. [16] designed a small laser terminal (SLT) using a QD as a spot-detector device, with a maximum communication distance of 3 km and bit-error rate (BER) of 10^{-6}. The basic block diagram of the SLT is shown in Fig. 9.8.

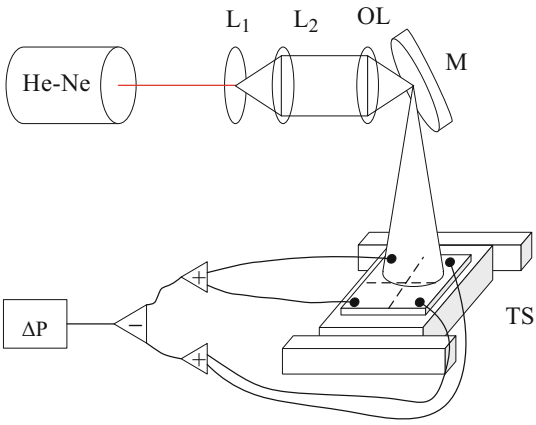


Fig. 9.7 Experimental-setup diagram [15]

The optical system and objective lens split the incident and outgoing beams of the received laser signal between the data receiver and the four-quadrant detector. The digital signal processor controls the actuator, according to the four-quadrant detector. The beam-coupling device is responsible for separating the input and output laser beams. The output laser beam is generated by the laser module and amplified by the optical amplifier.

In 2013, Barbarić et al. [17] derived a new relationship between the QD area and position to analyze a laser-tracker system model. The results showed that the new relation signal processing could improve the tracking accuracy by 30%.

In 2013, the Lunar Atmosphere and Dust Environment Explorer (LADEE) carried the optical communication Lunar Lasercom Space Terminal (LLST) for two-way laser-communication experiments with the ground [18–20]. The APT subsystem is shown in Fig. 9.9. The telescope is mounted on a gimbal to ensure the coarse alignment of the optical link over a wide range. The rate sensor is used to detect angular perturbations, which are then suppressed by a voice coil actuator. The QD receives the beam, determines the spot position, and drives a fiber-optic chapter actuator to achieve fine tracking.

In 2021, Safi et al. [21] investigated a ground-to-air free-space optical link for a multirotor hovering unmanned aerial vehicle (UAV), using a four-quadrant array of photodetectors on an optical receiver to detect light spots.

Domestic Research Progress

Chinese researchers and scholars have also conducted in-depth research on QD-based laser-spot position-detection technology, studied the factors affecting the spot-position detection accuracy, proposed different detection algorithms, and designed different QD spot-detection systems. The research progress is as follows.

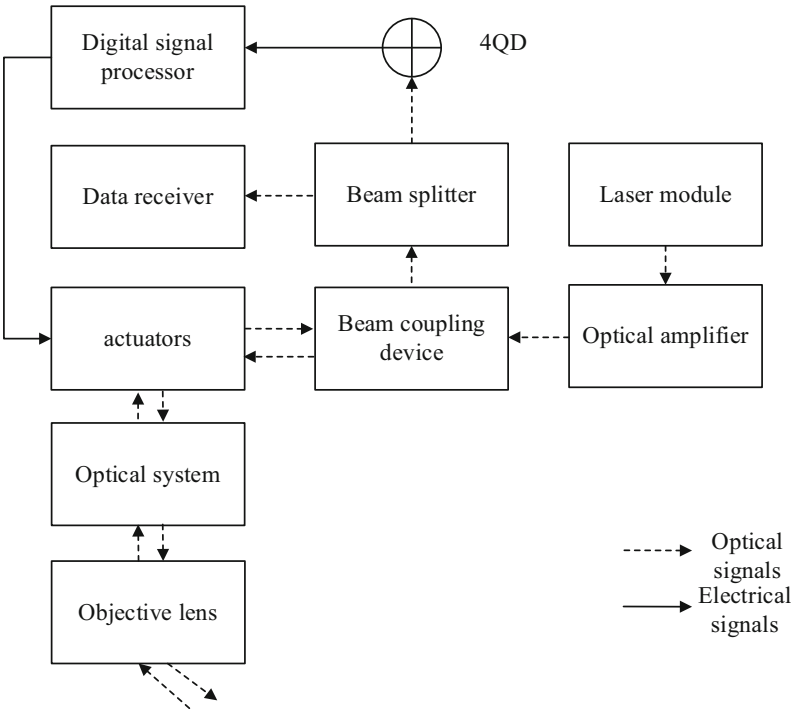


Fig. 9.8 Basic block diagram of SLT [16]

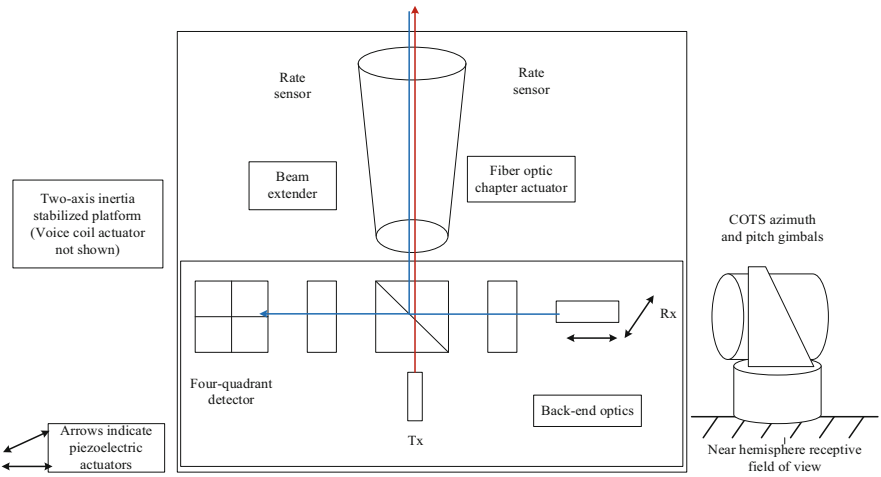


Fig. 9.9 LLST optical-path diagram [46]

In 1996, Feng Longling et al. [45] conducted a more detailed computational analysis of the error-data extraction algorithm, spot size, and other influencing factors in QD tracking technology.

In 1996, Tian et al. [42] described the principle of the sum-difference detection algorithm for angle measurement, carried out error and noise analyses, and found that spot and atmospheric-turbulence noise were the main error sources.

In 1997, Zhou [44] designed an electromechanical four-quadrant detector to measure the spot position. Its measurement resolution was better than 0.1 μm.

In 1999, Wan et al. [43] designed a beam fine-tracking system using a QD and fast reflector and tested the system performance. The results showed that the system had a good fine-tracking capability.

In 2003, Wang et al. [22] designed a two-axis tracking-control experimental demonstration system using QD to measure the spot position and digital signal processing (DSP) to control fast and wide-range scanning systems that could quickly capture and smoothly track the target.

In 2007, Xu [23] studied the relationship between the QD error signal and the spot position, size, and detector and analyzed and quantitatively estimated the causative factors. They proposed an optimized spot-size design and concluded that a spot size slightly larger than one-half of the effective radius of the detector could help improve the system performance.

In 2009, Chen et al. [9] proposed an improved algorithm combining the interpolation method and a diagonal algorithm. It first calculated the spot center using the actual offset and then processed the results of the traditional algorithm, using the interpolation method to bring it closer to the actual value. The angular error could be controlled to within 0.1°.

In 2010, Zhao et al. [5, 24] studied various QD properties and the influence of the external environment on their performance. They conducted experimental tests to examine the ultimate sensitivity, subdivision ability, signal-to-noise ratio, background light, spot size, and other characteristics.

In 2012, Chen et al. [25] discussed and compared four algorithms—sum-difference, diagonal, Δ/Σ, and logarithmic—for various light-spot models.

In 2012, Li et al. [26] proposed the gain-adjustable fast tracking and positioning (GAFT) algorithm. The GAFT algorithm improved the tracking accuracy to 0.0497 mm by changing parameters such as the spot radius and signal strength.

In 2013, Zhang et al. [27] developed a small-laser tracking system using a QD as a spot detector. The subsequent detector circuit had a simple structure and low complexity.

In 2017, Guo et al. [28] proposed and simplified a bipartite polynomial-fitting algorithm to improve the accuracy of QDs in detecting the spot-center position. The spot-detection accuracy of the QD was improved to the order of 10^{-4} mm.

In 2017, Li [6] proposed a spot-detection method to recognize the spot-absence state and adjust it to improve the detection accuracy and reduce the detection error.

In 2017, Zhang et al. [29] proposed a scheme for tracking and communication multiplexing using an avalanche diode–type QD. When the spot diameter was about half of the QD photosensitive-surface diameter, the minimum resolution for spot-position detection was 2 μm, the detection range could reach 0.8 mm, and the angular resolution was 0.8 μrad.

In 2018, Lin et al. [30] designed and tested a QD-based laser-beam 2D scanning–-tracking system. The test results showed that this tracking and scanning system could track moving targets more accurately and in real time.

In 2019, Liu et al. [31] proposed a QD-based laser-tracking system with a tracking error of approximately 0.1% and a tracking speed of 22.6°/s with a ± 20° moving change in the target position.

In 2019, Tong et al. [32] used a QD as a probe for a communication receiver to design a composite digital-tracking communication receiver. A 6.29-μm theoretical position resolution was obtained at a spot radius of 400 μm. A detection sensitivity of −30 dBm was obtained by experimental testing at a BER of 10^{-6} and a communication rate of 10 Mbit/s.

In 2020, Zou et al. [33] proposed a center opening-type four-quadrant detector fiber-positioning mechanism and designed a positioning algorithm. It calibrated the unit-spot bundle waist once and achieved multiple high-precision real-time spot-position determinations and fiber-position adjustments.

In 2021, Liang [34] discussed the relationship between the spot and QD output for different energy-distribution models and studied the center-positioning algorithm and spot-detection working mode for through-hole four-quadrant detectors.

In 2021, Liu et al. [35] proposed an improved algorithm (GII) based on the infinite integral-fitting method. The linear detection range of the GII algorithm was better than that of the traditional infinite integral-fitting algorithm and the eighth-order polynomial-fitting method, with high sensitivity.

In 2021, Wei et al. [36] improved the localization algorithm under a Gaussian spot model to improve the localization accuracy and linear range. It was experimentally verified that the root-mean-square error of the improved Gaussian algorithm was reduced by 60.4%, compared with the conventional Gaussian algorithm.

In 2021, Li et al. [37] investigated the effect of each noise factor when using QD for speckle-position localization. They found that increasing the transmit power could improve the signal-to-noise ratio of the signal at the receiver end, and improve the angular accuracy within a certain range.

In 2022, Xu and Long et al. [38] established the characterization equations of the QD output versus the spot position, considering the spot radius, energy distribution, and detector dead zone. In the 300-μm range, the QD spot-center localization accuracy was 0.35 μm and the root-mean-square error was 0.18 μm.

In 2022, Chen et al. [39] designed a set of micronano laser-communication terminals based on a QD and used a micro-electromechanical system (MEMS)

oscillator as the servo architecture to detect and track an optical spot. The experimental results showed a tracking error of 84 μrad.

QD Spot-Detection Influence Factors

The factors affecting the spot-position detection accuracy are the spot shape and distribution, QD defects, environmental factors, optical and electrical noise, and the detection algorithm.

Effect of Different Light-Spot Modes on the QD Output

Different spot patterns affect the output of the QD, and the generally studied energy-distribution models are uniform and Gaussian. From Eq. (9.12), it can be observed that for a Gaussian spot, the output of the QD in the x-direction with respect to the spot-moving distance can be expressed as [38]

$$\sigma_x = \frac{(S_A + S_D) - (S_B + S_C)}{S_A + S_B + S_C + S_D} = 1 - \left(\frac{2\theta}{\pi} - \frac{2x_0 \sin\theta}{\pi r}\right), \sigma_x = erf\left(\frac{\sqrt{2}x_0}{r}\right), \quad (9.14)$$

where $\theta = acr\cos\left(\frac{x_0}{r}\right)$, σ_x is the actual spot offset, r is the spot radius, x_0 is the moving distance, and S_A, S_B, S_C, and S_D denote the spot areas incident to the four quadrants, respectively.

Figure 9.8 in Reference [37] shows the relationship between the spot center and the QD output value for different spot modes. The spot-position detection sensitivity of a uniform spot is poor, and the linear dynamic range is large. The spot-position detection sensitivity of a Gaussian spot is high, and the linear dynamic range is small.

Effect of Different Spot Radii on the QD Output

Assuming a uniform spot model, the relationship between the spot center and QD output value is investigated for different spot radii. See Fig. 9.4 in Reference [37], with spot radii of $r = 1$, 2, and 3 mm. As the spot radius increases, the detection sensitivity of the QD decreases.

Effect of Different Dead-Zone Widths on the QD Output

A dead zone exists between every two adjacent QD quadrants. The dead-zone size affects the detector's reception of the total spot energy and the accuracy of the spot

localization. When the dead-zone width is d, the detector output in the x direction with respect to the spot-moving distance can be expressed as [38]

$$\sigma_x = 1 - \left(2\theta r^2 - 2x_0 r\sin\theta - \left(2d(r - x_0) + 2d\sqrt{r^2 - x_0^2}\right)\right) / \left(\pi r^2 - \left(2dr + 2d\sqrt{r^2 - x_0^2}\right) - d^2\right). \tag{9.15}$$

Assuming a 2-mm spot radius and dead-zone widths of 0, 0.04, 0.2, and 0.3 mm, Fig. 9.6 in Reference [37] shows that the linear dynamic range of the QD decreases as the dead-zone width increases.

Effect of Different Background Lights on the QD Output

The presence of background light increases the total energy received by the QD, thus increasing the output current. Usually, the ambient-light intensity can be considered uniform, and thus, the background-light distribution in each quadrant on the QD can also be considered uniform. Thus, the output signal of the QD in the x direction with respect to the spot-moving distance can be expressed as [48]

$$\sigma_x = \frac{(I_A + I_D) - (I_B + I_C)}{I_A + I_D + I_B + I_C + I_{BF} + I_{DK}}, \tag{9.16}$$

where I_A, I_B, I_C, and I_D represent the photocurrents generated in each quadrant, I_{BF} represents the photocurrent generated by the entire QD illuminated by background light, and I_{DK} represents the dark current of the entire QD.

Figure 9.3 in Reference [48] shows the relationship between the spot center and detector output for different ratios of background light. As the background light increased, the spot-position detection accuracy and sensitivity decreased to some extent.

Effect of Different Signal-to-Noise Ratios on the QD Output

Improving the system signal-to-noise ratio can improve the position-detection accuracy of a QD. The relationship between the standard deviation of the spot position, influenced by the spot radius, position, and system S/N ratio, is [46]

$$\sigma_r = \frac{\sqrt{\pi}r}{2\sqrt{2}} \frac{1}{\sqrt{SNR_{\sum}}} \left[1 + \frac{x^2 erf^2\left(\frac{\sqrt{2}x}{r}\right) + y^2 erf^2\left(\frac{\sqrt{2}y}{r}\right)}{x^2 + y^2}\right], \tag{9.17}$$

where SNR_{Σ} is the total signal-to-noise ratio of the system, r is the spot radius, and x and y are the spot positions.

Figure 9.4.4 in Reference [46] shows that as the signal-to-noise ratio increases, the standard deviation of the spot position decreases.

Effect of Different Positioning Algorithms on the QD Output

The spot center-of-mass detection algorithm is an important factor that affects the accuracy of the QD detection. The commonly used algorithms are the addition-subtraction, diagonal, difference-ratio-sum, and logarithmic algorithms. Assuming that the spot is a uniform circular spot, the four algorithms are compared; see Fig. 9.6 in Reference [35].

It can be seen that for a circular spot with a uniform energy distribution, the linearity and sensitivity of the four algorithms differ. The difference-ratio-sum algorithm is the best in terms of sensitivity and the worst in linear range. The diagonal algorithm has good linearity but is lower than the difference-ratio-sum algorithm and has reduced sensitivity. The addition-subtraction algorithm is more balanced in sensitivity and linear range, and the logarithmic algorithm has a wider linear range.

The missing spot was modeled as an ideal elliptical spot, and its impact on the spot detection was simulated and analyzed. The offsets Δx, Δy, Δx_T, and Δy_T in the x- and y-axis directions for circular and elliptical spots are

$$\Delta x = P_0 erf\left(\frac{\sqrt{2}x_0}{R}\right) \tag{9.18}$$

$$\Delta y = P_0 erf\left(\frac{\sqrt{2}y_0}{R}\right) \tag{9.19}$$

$$\Delta x_T = P_0 erf\left(\frac{\sqrt{2}x_0}{w_a}\right) \tag{9.20}$$

$$\Delta y_T = P_0 erf\left(\frac{\sqrt{2}y_0}{w_b}\right), \tag{9.21}$$

where P_0 is the received optical power, x_0 is the travel distance in the x direction, y_0 is the travel distance in the y direction, r is the radius of the circular spot, w_a is the long axis of the elliptical spot, and w_b is the short axis of the elliptical spot.

Figure 9.6 in Reference [6] shows the variation curves of the x- and y-axis offsets with offset detection values under different spot models. It can be observed that the larger the spot variation, the smaller the linear interval. The missing spot reduces the detectable range of the QD.

Spot-Detection Algorithms

When a QD is used to detect the spot position, the spot center-of-mass position is determined according to the four signals output from the detector. Commonly used algorithms include addition-subtraction, diagonal, difference-ratio-sum, and logarithmic. Function-fitting methods include normalized-center, polynomial-fitting, infinite-integration, Boltzmann function-fitting, composite, and infinite integration-improvement.

Addition-Subtraction Algorithm

When a spot is irradiated onto the QD, the offset of the spot relative to the center of the detection surface is denoted by σ_x and σ_y, and the offset is calculated using the addition and subtraction algorithm to obtain [8]

$$\sigma_x = \frac{(U_A + U_D) - (U_B + U_C)}{U_A + U_B + U_C + U_D} \tag{9.22}$$

$$\sigma_y = \frac{(U_A + U_B) - (U_C + U_D)}{U_A + U_B + U_C + U_D}, \tag{9.23}$$

where U_A, U_B, U_C, and U_D denote the voltages in the four quadrants of the QD.

For the QD, the x- and y-axis offsets are proportional to the spot area, so the actual offsets σ_x and σ_y can be expressed as

$$\sigma_x = \frac{(S_A + S_D) - (S_B + S_C)}{S_A + S_B + S_C + S_D} \tag{9.24}$$

$$\sigma_y = \frac{(S_A + S_B) - (S_C + S_D)}{S_A + S_B + S_C + S_D}, \tag{9.25}$$

where S_A, S_B, S_C, and S_D are the spot areas in the four QD quadrants.

Diagonal Algorithm

The diagonal algorithm was generated to extend the linear region of the measurement. Using the diagonal algorithm to calculate offsets σ_x and σ_y yields

$$\sigma_x = \frac{U_A - U_C}{U_A + U_B + U_C + U_D} \tag{9.26}$$

$$\sigma_y = \frac{U_B - U_D}{U_A + U_B + U_C + U_D}. \tag{9.27}$$

The relationship between the spot-center offset and spot area in each quadrant of the detector in the linear region is

$$\sigma_x = \frac{S_A - S_C}{S_A + S_B + S_C + S_D} \tag{9.28}$$

$$\sigma_y = \frac{S_B - S_D}{S_A + S_B + S_C + S_D}. \tag{9.29}$$

Difference-Ratio-Sum Algorithm

The measurement sensitivity of the diagonal algorithm is reduced compared to that of the addition-subtraction algorithm, and the difference-ratio-sum algorithm was proposed to improve the measurement sensitivity. Using the difference-ratio-sum algorithm to calculate the offsets σ_x and σ_y yields

$$\sigma_x = \frac{U_A - U_C}{U_A + U_C} \tag{9.30}$$

$$\sigma_y = \frac{U_B - U_D}{U_B + U_D}. \tag{9.31}$$

In the linear region, the relationship between the spot-center offset and spot area in each quadrant of the detector is

$$\sigma_x = \frac{S_A - S_C}{S_A + S_C} \tag{9.32}$$

$$\sigma_y = \frac{S_B - S_D}{S_B + S_D}. \tag{9.33}$$

Logarithmic Algorithm

The logarithmic algorithm has a high bandwidth, wide dynamic range, and good linearity. Using the logarithmic algorithm to calculate the offsets σ_x and σ_y [9],

$$\sigma_x = \log\left(\frac{U_A}{U_C}\right) \tag{9.34}$$

$$\sigma_y = \log\left(\frac{U_B}{U_D}\right). \tag{9.35}$$

In the linear region, the relationship between the spot-center offset and spot area in each quadrant of the detector is

$$\sigma_x = \log\left(\frac{S_A}{S_C}\right) \tag{9.36}$$

$$\sigma_y = \log\left(\frac{S_B}{S_D}\right). \tag{9.37}$$

Function-Fitting Methods

The function-fitting method [45] uses various types of functions to fit the relationship between the solved QD value and the actual position of the spot center to determine the position of the spot center. Their relationship curve is an "S" type.

(1) **Normalized Center**

The normalized-center method approximates the response within a certain region near the center of the QD as a linear response and fits it with a first-order function to obtain the actual location of the spot center [47]:

$$x_0 \approx k \times \sigma_x, \tag{9.38}$$

where k is the scale factor, which is related to the QD shape, size, spot-intensity distribution, and dead-zone width. The normalized-center method error increases as the spot center gradually moves away from the detector center.

(2) **Polynomial Fitting**

The polynomial-fitting method uses a higher-order polynomial to fit a function between the solution value and the actual location of the spot center [45]:

$$x_0 \approx \sum_{i=0}^{n} a_n \sigma_X^n, \tag{9.39}$$

where a_n is the fitting coefficient. The accuracy of the polynomial-fitting method depends on the order; the higher the order, the higher the accuracy, but the higher the hardware requirements.

(3) **Infinite Integration**

The infinite-integration method assumes that the QD has no dead zone and is infinite. Its solution value can be expressed as [48]

$$\begin{aligned}\sigma_x &\approx \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{0} I(x,y)dxdy - \int_{-\infty}^{\infty}\int_{0}^{\infty} I(x,y)dxdy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(x,y)dxdy} \\ &= \frac{\frac{I^2\pi}{4}\left(1+erf\left(\frac{\sqrt{2}}{r}x_o\right)\right) - \frac{I^2\pi}{4}\left(1-erf\left(\frac{\sqrt{2}}{r}x_o\right)\right)}{\frac{I^2\pi}{2}}, \\ &= erf\left(\frac{\sqrt{2}x_0}{r}\right)\end{aligned} \tag{9.40}$$

where $erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the error function. The actual position of the spot center is obtained by inverting both sides of Eq. (9.40):

$$x_0 \approx \frac{erf^{-1}(\sigma_x)r}{\sqrt{2}}. \tag{9.41}$$

(4) **Boltzmann-Function Fitting**

Boltzmann-function fitting involves finding the functional relationship between the solved value and the actual position of the spot center. The functional expression is [46]

$$y = \frac{A_1 - A_2}{1 + e^{(x_0 - x_a)/k}} + A_2, \tag{9.42}$$

where A_1, A_2, x_a, and k are parameters to be determined. By performing the following mappings, $y = \sigma_x$, $A_1 = -1$, $A_2 = 1$, and $x_a = 0$, we obtain [46]

$$\sigma_x \approx 1 - \frac{2}{1 + e^{x_0/k}}, \tag{9.43}$$

where k is the scale factor, which is related to the QD shape, size, spot-intensity distribution, and dead-zone width. The expression of the Boltzmann-function fitting method is [46]

$$x_0 \approx k \times \ln\left(\frac{1+\sigma_x}{1-\sigma_x}\right). \tag{9.44}$$

(5) **Composite Fitting**

Composite fitting is a new fitting algorithm that linearly fuses two algorithms: Boltzmann-function fitting and infinite fitting. Its spot center-of-mass position is expressed as [47]

$$x_0 \approx m \times k_1 \frac{erf^{-1}(\sigma_x) \cdot \omega}{\sqrt{2}} + (1-m) \times k_2 \ln\left(\frac{1+\sigma_x}{1-\sigma_x}\right), \tag{9.45}$$

where k_1 and k_2 are correction coefficients and m represents the weights. k_1, k_2, and m can be obtained by least squares using N sets of experimental or simulation data [46]:

$$k_1 = \frac{\sum_{n=1}^{N} f_1(\sigma_{X_n}) \times x_n}{\sum_{n=1}^{N} f_1^2(\sigma_{X_n})}, \quad f_1(E_X) = \frac{erf^{-1}(\sigma_X)}{\sqrt{2}} \tag{9.46}$$

$$k_2 = \frac{\sum_{n=1}^{N} f_2(\sigma_{X_n}) \times x_n}{\sum_{n=1}^{N} f_2^2(\sigma_{X_n})}, \quad f_2(E_X) = \ln\left(\frac{1+\sigma_X}{1-\sigma_X}\right) \tag{9.47}$$

$$m = \frac{\sum_{i=1}^{N} (x_n - k_1 f_1(\sigma_{X_n}))(k_1 f_1(\sigma_{X_n}) - k_2 f_2(\sigma_{X_n}))}{\sum_{i=1}^{N} (k_1 f_1(\sigma_{X_n}) - k_2 f_2(\sigma_{X_n}))^2}. \tag{9.48}$$

(6) **Infinite Integral-Fitting Improvement**

The infinite integral-fitting improvement algorithm is proposed to improve the positioning accuracy by considering the effects of the QD photosensitive-surface radius and dead-zone width. The compensation function is fitted by introducing a compensation factor and using a segmented polynomial-fitting method. The expression for the spot center-of-mass position is [46]

$$x_0 = \frac{erf^{-1}(\sigma_x)}{\sqrt{2}} \times r \times \lambda(r, R, d), \tag{9.49}$$

where r is the Gaussian spot radius, R is the QD radius, and d is the dead-zone width.

By combining the last two terms of Eq. (9.49) into a single parameter, it becomes [46]

$$x_0 = g(\sigma_x) \times \lambda_e(\sigma_x, r, R, d). \tag{9.50}$$

The least-squares method was used to model the position-detection error [46]:

$$I(\lambda_e) = \|\delta_x\|^2 = \sum_{i=1}^{N} [x_0(\lambda_e, \sigma_{x_i}) - x_i]^2. \tag{9.51}$$

Deriving the equation gives [46]

$$\frac{\partial I}{\partial \lambda_e} = 2\sum_{i=1}^{N} g(\sigma_x) \times \lambda_e - 2\sum_{i=1}^{N} g(\sigma_x) \times X_i. \tag{9.52}$$

Let $\frac{\partial I}{\partial \lambda_e} = 0$, to obtain the optimal λ_e [46]:

$$\lambda_e = \frac{\sum_{i=1}^{N} g(\sigma_x) \times X_i}{\sum_{i=1}^{N} g^2(\sigma_x)}. \tag{9.53}$$

QD Light-Spot Detection System

Beam-Alignment Detection System

The structure of the beam-alignment detection system is shown in Fig. 9.10. The optical receiving antenna uses a Cassegrain telescope. A detection module is placed behind the receiving telescope, and a cylindrical holder, dimming lens, and QD are placed vertically above the beam-splitting prism. The system detects the beam alignment using two motors.

Missing-Spot Detection Test

In Reference [40], Table 9.2 shows the actual position-measurement data obtained with the incident beam in the missing state. The long axis w_a of the tested spot was 0.6 mm, the short axis w_b was 1.2 mm, and the right side of the spot was missing. It can be observed that the actual value of the detector offset deviates from the calculated value under the missing spot. The detected missing spot is shown in Fig. 9.11.

System Alignment Test

Figure 9.12 shows the complete spot detected when the spot is adjusted to the no-detect state using the threshold-detection algorithm.

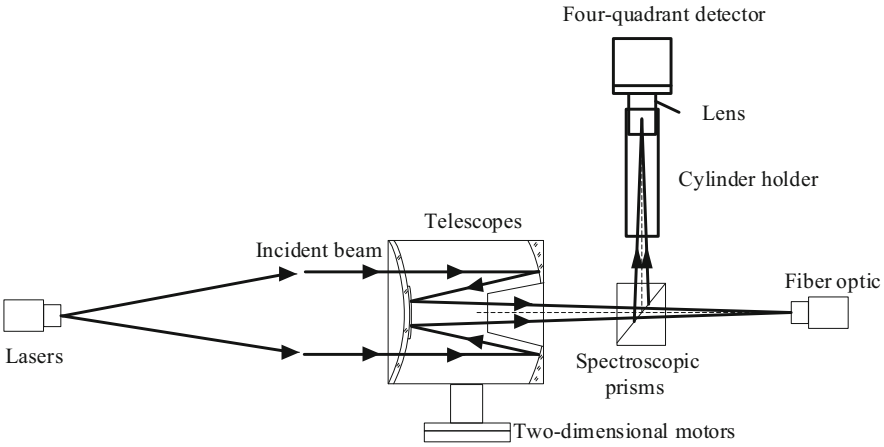


Fig. 9.10 System architecture [41]

Fig. 9.11 Missing-spot detection map [41]

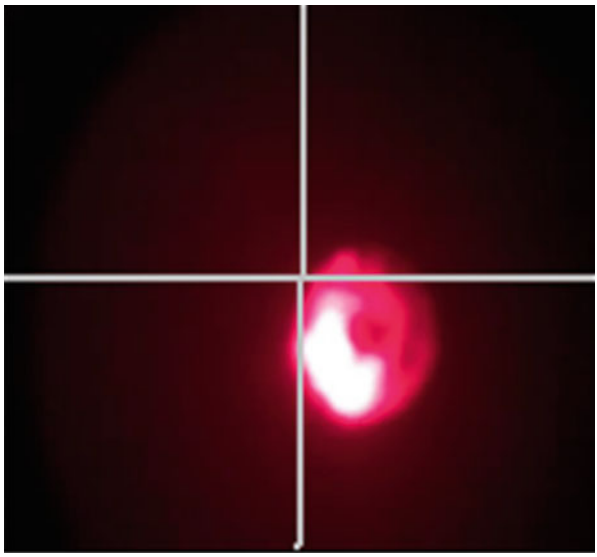


Figure 9.13 shows the output of the detector when the spot is adjusted from the missing state to the full-spot state.

QD Composite Tracking and Communication System

The QD composite tracking and communication detection system [29] is shown in Fig. 9.14 and consists of four parts: a laser-modulated emission unit, signal detection and processing unit, tracking control unit, and communication unit.

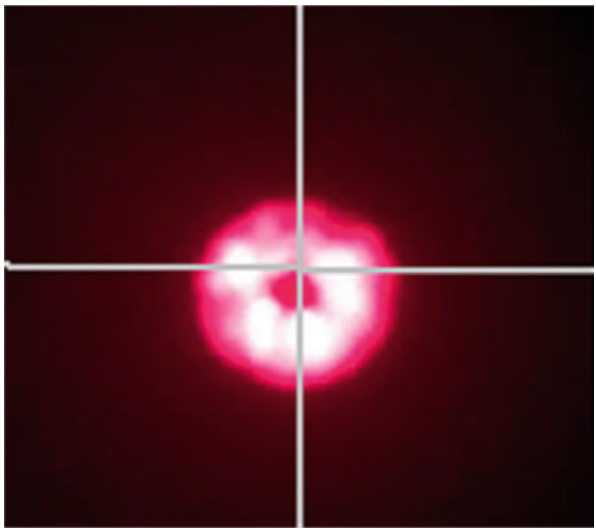


Fig. 9.12 Complete-spot detection map [41]

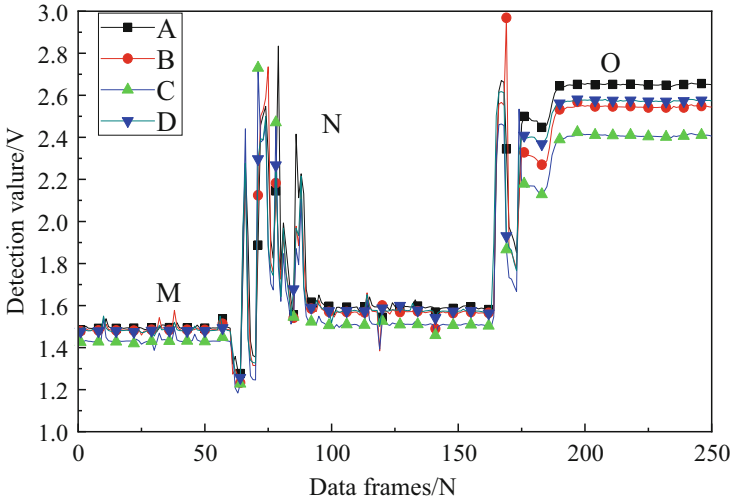


Fig. 9.13 QD output diagram of the regulation process [41]

Micro- and Nano-laser-communication Terminal

The optical architecture and servo-system structure of the micro- and nano-laser-communication terminal [39] is shown in Fig. 9.15. It consists of a four-quadrant detector, MEMS oscillator, oscillator driver, analog-to-digital converter, and digital-to-analog converter. The light spot is tracked using the four-quadrant detector and MEMS oscillator.

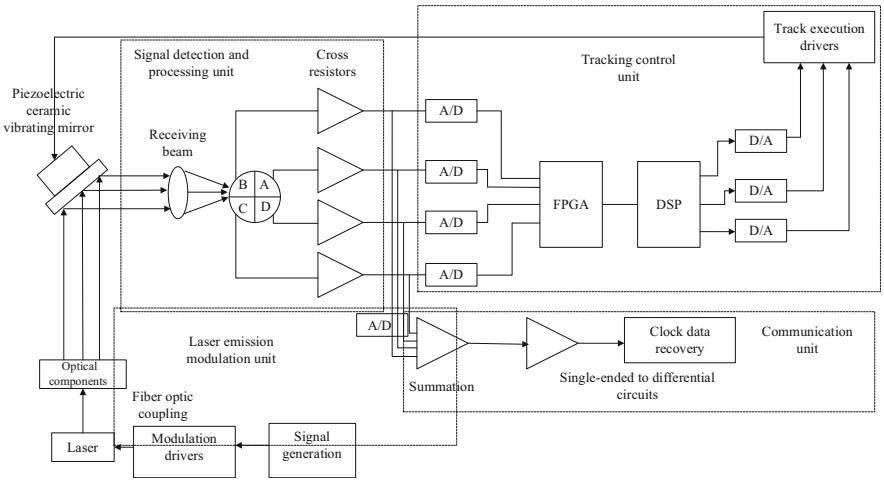


Fig. 9.14 Composite tracking and communication detection system [29]

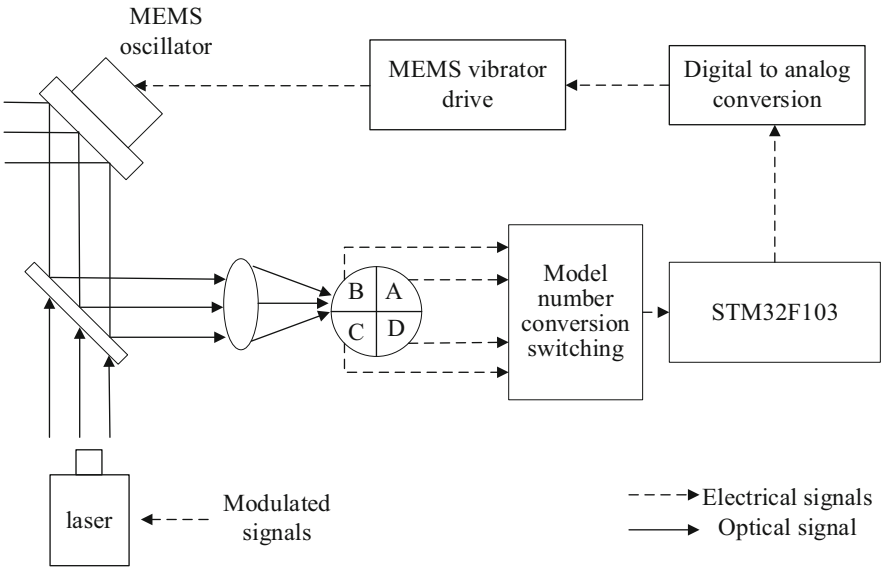


Fig. 9.15 Schematic diagram of the micro- and nano-laser-communication terminal system [39]

Summary

This chapter used spot detection as the entry point, introduced the basic spot model and QD detection principle, and reviewed the QD spot-detection development status. In addition, it described the QD spot-detection influencing factors and spot-detection

algorithms and introduced the QD spot-detection system. Spot-detection technology based on the four-quadrant detector is likely to develop in the following directions.

1. Use APD-type QDs to achieve composite tracking and communication detection, integrate tracking and communication, miniaturize servo-structure design, and reduce the size and power consumption of the laser-communication system.
2. Study the spot-position detection algorithm, improve the spot-position detection accuracy, and reduce the requirements for hardware processing capability.
3. Establish a complete mathematical model for describing and analyzing the effect of noise on the spot-position detection accuracy and study in depth how to suppress the noise.
4. Study the effects of different combinations of influencing factors on the detector accuracy.

References

1. Li Xi. Research on QD-based Precise Positioning Technology [D]. Xi'an, Shanxi: Xi'an University of Electronic Science and Technology, 2012.
2. Li D, Liu S, Bin W. Research on Factors that Influence Detecting Precision of Quadrant Detector [J]. International Journal of Advancements in Computing Technology,2011, 3(2): 112–118.
3. Hermosa N, Aiello A, Woerdman J P. Quadrant detector calibration for vortex beams [J]. Optics Letters, 2011, 36(3):409–412.
4. Meng Y, Sun A X, Zhang G Y, et al. Research Effect of Four-Quadrant Detector Detection Spot in Atmospheric Turbulence [J]. Applied Mechanics & Materials, 2014, 552(2):123–128.
5. Zhao Xin, Tong Shoufeng, Jiang Huilin. Characterization of four-quadrant detectors [J]. Optical Precision Engineering,2010,18(10):2164–2170.
6. Ke Xizheng, Li Shiyan. Experimental study of optical antenna optical axis alignment in the case of spot fragmentation [J]. Journal of Photonics,2017,46(04):133–141.
7. Guo H W, Liu R Q, Deng Z Q, et al. Performance Analysis and Testing of Four-Quadrant Position Sensitive Detector [J]. Advanced Materials Research, 2011, 317(3):1107–1114.
8. Toyoda M, Araki K, Suzuki Y. Measurement of the characteristics of a quandrant avalanche photo-diode application to a laser tracking system [J]. Optical Engineering, 2002, 41(1): 145–1410.
9. Chen Mengwei. Research on light spot center localization algorithm based on image detector [D]. Wuhan, Hubei, China: Wuhan University of Technology, 2013.
10. H. Hemmati, A. Biswas, I. B. Djordjevic, et al. Deep-Space Optical Communications: Future Perspectives and Applications [J]. Proc. of the IEEE, 2011, 99(11): 2020–20310.
11. Takashi J. Optical Inter-orbit Communication Experiment between OICETS and ARTEMIS [J]. Journal of the National Institute of Information and Communications Technology, 2012, 59(1–2):23–33.
12. Bailly M. The pointing, acquisition and tracking system of SILEX European program – A major technological step for intersatellites optical communication [J]. SPIE, 1991, 1417(5):142–158.
13. Guelman M, Livne A, Michalik H, et al. Broadband laser intersatellite link for microsatellites: project overview [J]. Space technology, 2002, 22(3):113–1110.
14. Toyoshima M, Takenaka H, Shoji Y, et al. Results of Kirari optical communication demonstration experiments with NICT optical ground station (KODEN) aiming for future classical and quantum communications in space [J]. Acta astronautica, 2012, 74(1):40–410.

15. Lee E J, Park Y, Kim C S, et al. Detection sensitivity of the optical beam deflection method characterized with the optical spot size on the detector [J]. Current Applied Physics, 2010, 10(3):834–838.
16. Schmidt C, Horwath J. Wide-field-of-regard pointing, acquisition and tracking-system for small laser communication terminals [J]. Proceedings of ICSOS 2012,8(4):1–6.
17. Žarko P. Barbarić, Stojadin M. Manojlović, Boban P. Bondžulić, et al. New relationship of displacement signal at quadrant photodiode: Control signal analysis and simulation of a laser tracker [J]. Optik-International Journal for Light and Electron Optics, 2013, 125(4): 1550–1558.
18. Z. Chen, J. Yan. Impact of pointing errors on performance of a ground-to-satellite laser uplink communication system based on M-ary pulse position modulation [J]. The Journal of China Universities of Posts and Telecommunications, 2009, 16(Supplement 1):20–23.
19. B. S. Robinson, D. M. Boroson, D. A. Burisnrk, et al. The Lunar Laser Communications Demonstration [C]. IEEE 2011 International Conference on Space Optical Systems and Applications, 2011,7923(1):54–57
20. B. S. Robinson, D. M. Boroson, D. A. Burianek, et al. Overview of the Lunar Laser Communications Demonstration [J]. Proc. of SPIE, 2011, 8610(10):02.
21. Safi H, Dargahi A, Cheng J. Beam Tracking for UAV-Assisted FSO Links With a Four-Quadrant Detector [J]. 2021,25(12):3908–3913.
22. Wang D., Yang S. H. Two-axis tracking control technique using four-quadrant detector [J]. Optical Engineering,2003,34(05):31–33+48.
23. Xu Daisheng. Optimal design of signal spot for four-quadrant detection system [J]. Journal of Hunan Institute of Technology (Natural Science Edition),2007,20(01):50–53+90.
24. Zhao X, Tong Shoufeng, Liu Yunqing, et al. Research on the application of four-quadrant detectors in space laser communication [J]. Optoelectronics. Laser, 2010,21(01):46–410.
25. Chen Mengwei, Yang Yingping, Jia Xinting, et al. A four-quadrant detector spot centering algorithm [J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition), 2013,37(05):1124–1128.
26. Si Dongsen, Li Zengzhi, Wang Xiaoxu. Intelligent tracking and positioning algorithm using four-quadrant detector [J]. Journal of Xi'an Jiaotong University,2012,46(04):13–17+910.
27. Zhang W, Zhang H, Liu B, et al. Design of Small Laser Tracking System Based on Four-Quadrant Detector [J]. Applied Mechanics & Materials, 2013, 30(1):393–3910.
28. Guo S. K., Zhang Y. M., He S. J.. High-precision localization algorithm for spot center based on four-quadrant detector [J]. Laser and Infrared,2017,47(11):1353–1358.
29. Fan XK, Zhang L, Song YNS, et al. Tracking and communication composite detection technology for four-quadrant detectors [J]. China Laser,2017,44(09):247–254.
30. Lin Xin, Guo Ying, Han Mingzhu, et al. Research on Two Dimensional Laser Beam Scanning and Tracking System Based on Four Quadrant Detector [J]. Semiconductor Optoelectronics, 2018, 39(03):425–430.
31. Liu S. M., He N., Deng D. Y.. Design and implementation of laser tracking system based on four-quadrant detection [J]. Journal of Guilin University of Electronic Science and Technology,2019,39(01):7–12.
32. Wang Ruiyang, Yu Xiaonan, Tong Shoufeng, et al. Design of a four-quadrant detector digital tracking communication composite receiver [J]. Journal of Changchun University of Science and Technology: Natural Science Edition, 2019,42(3):10.
33. Wang Linzheng, Zou Hua, Huang Shuo, et al. Closed-loop control method for fiber positioning of center-opening four-quadrant detectors [J]. Infrared and Laser Engineering,2020,49(06): 264–270.
34. Ke Xizheng, Liang Hanli. Airborne Laser Communication System with Automated Tracking [J]. International Journal of Optics, 2021,2021(5):1–9.
35. Yepeng Gou, Xing Liu, Qiang Liu, et al. Infinite integral-based four-quadrant detector spot position detection algorithm [J]. Laser and Infrared, 2021,51(9):1249–1256.
36. Qin Licun, He Wei. Laser spot center localization algorithm based on four-quadrant detector [J]. Applied Lasers,2021,41(3):561–569.

37. Li Shude, Liu Caixia, Xu Lin, et al. Positioning accuracy of four-quadrant detector in interplanetary laser communication [J]. Optical Communication Research,2021,47(04):61–65.
38. Diao Kuan, Liu Xiaojun, Xu Long, et al. Study on the factors affecting the spot positioning accuracy of four-quadrant detectors [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition),2022,50(04):14–19.
39. Chen Yun, Yu Xiaonan, Jiang Lun, et al. Research on servo technology of micro-nano laser communication terminal based on QD and MEMS vibrating mirror [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2022,45(01):39–44.
40. Fang Xulai. Research on the application of four-quadrant detector in space laser communication [D]. Changchun, Jilin: Changchun University of Technology, 2018.
41. Li Shiyan. Research on experimental scheme for precise spot alignment in wireless laser communication system [D]. Xi'an, Shaanxi: Xi'an University of Technology, 2018.
42. Wan M, Zhang J. Ru, Nie W. J., et al. Fast tilting reflector precision tracking system [J]. Laser and Infrared, 1999,29(03):17–110.
43. Zhou M.P.. Double-knife combination for measuring microspot center position [J]. Optical Engineering, 1997,24(S1):93–98.
44. Feng Longling, Deng Renliang. A discussion of some problems in four-quadrant photoelectric tracking technology [J]. Infrared and Laser Engineering, 1996,25(01):16–21+18.
45. Wu Jia Bin. Research on key technology of high-precision laser spot position detection based on four-quadrant detector [D]. Beijing: Chinese Academy of Sciences, 2016.
46. Wu Jiabin, Chen Yunshan, Gao Shijie, et al. Improved measurement accuracy of spot position on an InGaAs quadrant detector [J]. Applied Optics, 2015,54(27):8049–8054.
47. Li Qian. Space laser communication spot position detection technology based on array detector [D]. Beijing: Chinese Academy of Sciences, 2020.
48. Zhang Jun, Qian Weixian, Liu Zewei. Analysis and correction of non-uniformity of four-quadrant detector output [J]. Infrared Technology, 2016,38(07):565–570.

Optical-Antenna Technologies

10

Contents

Introduction 298
Basic Structure of an Optical Antenna 298
Cassegrain Optical Antennas 302
- Large-Aperture Cassegrain Antenna 302
- Cassegrain Optical-Antenna Array 303
- Freeform-Surface Cassegrain Antenna 303
- Cassegrain Optical Antenna Based on a Beam-Shaping Scheme 304
- Off-Axis Dual-Reflection Cassegrain Optical Antenna 306
- Maksutov–Cassegrain Optical Antenna with Off-Axis Emission of a Gaussian-Beam Array 307
- Pre-Collimated Cassegrain Optical Antenna 311

Beacon-Free Tracking 311
Point-to-Multipoint Communication Optical Antennas 313
Summary 316
References 317

Abstract

Optical antenna is used to transmit or collect optical signals in free space optical communication system, and the structure of optical antenna determines the detection performance of the system. This chapter introduces the basic structures of optical antenna, the research progress of Cassegrain optical antenna, pointing, acquisition, and tracking (PAT) system without independent beacon and the point-to-multipoint optical antenna in detail, the future research direction of optical antenna is prospected.

Keywords

Free space optical communication · Optical antenna technology · Cassegrain system · Optical system without independent beacon · Point-to-multipoint optical antenna

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_10

Introduction

An optical antenna is the optical-signal receiving and transmitting system and is a key component of optical wireless communication systems. It transmits optical signals at the transmitting end and receives and couples optical signals at the receiving end. The performance of the optical system directly affects the performance and communication quality of the entire optical wireless communication system. Therefore, research on optical systems is highly significant for improving the performance of optical wireless communication systems and meeting the increasing demand for speed and capacity in future communications.

In this chapter, the basic structure of the optical antenna is introduced, and the widely used Cassegrain optical system is analyzed and compared. This chapter introduces a beacon-free optical system and a point-to-multipoint communication optical system, which will be necessary for the development of optical-communication networks in the future.

Basic Structure of an Optical Antenna

Since the United States launched the "Star Wars" program in 1953, space laser communication technology has rapidly developed. Germany, France, the United Kingdom, Japan, and other countries have also successively conducted relevant research [1], such as studying the working wavelength of an optical antenna as the transceiver of a communication system, beacon optical capture and tracking, and antenna-structure form.

The common structural forms of optical transceiver antennas are refractive, reflective, and reentrant. A refractive optical system is usually suitable for optical wireless communication in the visible and near-infrared bands. The reflection system is applicable to all wavebands. The reentrant optical system is suitable for long-distance optical wireless communication under complex atmospheric conditions, owing to its aberration correction.

Reflective optical systems can have multiple focal points, so they can produce multiple relative apertures, field angles, and focal lengths. Newtonian, Gregory, and Cassegrain optical systems are commonly used, as shown in Fig. 10.1a–c, respectively.

Newtonian optical systems have a relatively large aperture and are often used in optical systems with large apertures. However, the manufacturing cost is high, and they produce a coma for off-axis lights. The Gregory system simultaneously eliminates spherical and chromatic aberrations; however, the manufacturing process has high requirements, and there are few practical applications.

The Cassegrain optical system [2] is a dual-reflection telescopic optical system that has the advantages of a short system length, small volume, light weight, compact structure, and other practical characteristics. Moreover, the system has no chromatic or spherical aberrations, and it can adjust the system aperture according to its actual needs. Because of these advantages and characteristics, the Cassegrain optical system is widely used in free-space optical communications [3].

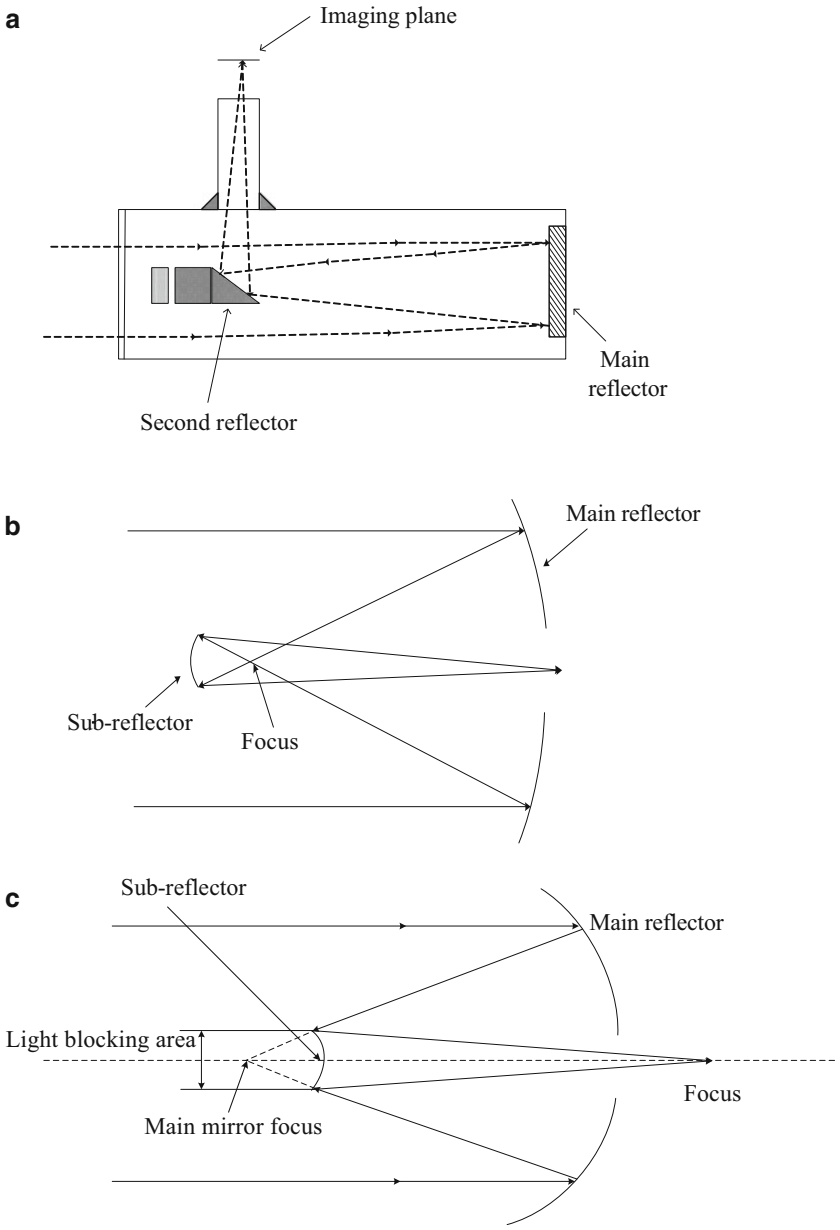


Fig. 10.1 Reflective optical systems. (**a**) Newtonian optical system, (**b**) Gregory optical system, (**c**) Cassegrain optical system

However, the occlusion ratio of the dual-reflection structure wastes optical-signal energy, serious off-axis aberrations exist, and high-quality image points are obtained at the expense of the field of view. Section "Beacon-Free Tracking" of this chapter

presents in-depth research on the defects and problems of the Cassegrain optical system.

The refractive optical system is shown in Fig. 10.2. The front objective is a group of glued lenses. After passing through the objective, the light is imaged in front of the eyepiece, and the image point is received after being magnified by the eyepiece. Refractive telescopes have the advantages of a wide field of view, high contrast, and clarity; however, the objective lens produces chromatic aberrations in refractive telescopes, and the larger the aperture, the more serious the chromatic aberration. Moreover, the lens-tube length determines the focal length, resulting in the large volume of the refracting telescope.

A structural diagram of the reentrant optical system is shown in Fig. 10.3. The reentrant optical system combines the objective lens of a refractive optical system with a reflective optical system. The objective lens is a wave-shaped correction lens with a convex center, concave periphery, and complex shape that corrects the spherical aberration of the spherical primary lens. The reentrant optical system mainly includes the Schmidt–Cassegrain, Maksutov–Cassegrain, and Fresnel optical systems.

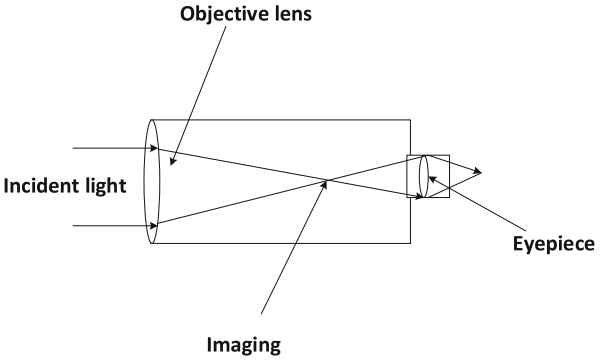


Fig. 10.2 Refractive optical system

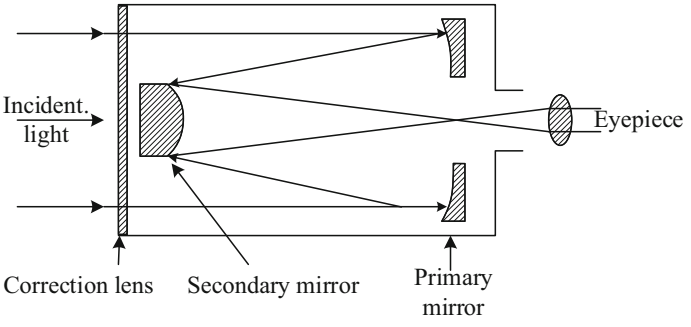


Fig. 10.3 Reentrant optical system

The Schmidt–Cassegrain system is based on the Schmidt astronomical observatory, uses a Schmidt correction lens to correct spherical aberrations, and adopts a convex mirror as the secondary mirror of the optical system, following the Cassegrain structure. It has many deformable structures, such as the double-spherical mirror group, double-aspheric lens group, and a combination of spherical and aspheric mirrors. This type of antenna structure is relatively simple to manufacture, compact in structure, and easy to carry.

A typical feature of the Maksutov–Cassegrain (Maka) system is that the secondary mirror is very small, usually an aluminized circular spot on the correction mirror. The catadioptric structure makes the Maka antenna compact, and a lens-group system with a large aperture and long focal length can be created with a small size. When the Maka antenna is used as a space optical-coupling system, it can produce large-aperture coupling receiving technology and has a strong ability to gather optical power. The long focal length easily meets the requirements of an optical-fiber numerical aperture. At the same relative aperture (D/f), the length of the lens system can be reduced to nearly one-half that of an ordinary lens system, and its advantages cannot be ignored in engineering.

The Fresnel optical system is based on the Cassegrain structure and uses a Fresnel lens to replace the refractive correction lens at the front end. Compared with conventional correction lenses, Fresnel lenses have the advantages of a strong focusing ability, short focal length, thin lens, and light weight. Their optical efficiency is more than 90%, which is suitable for most front-ends of optical systems with a small field of view and high gain [4]. The advantages and disadvantages of these three structures are compared in Table 10.1.

Table 10.1 Comparison of the refractive, reflective, and reentrant structures

Optical-system type	Advantages	Disadvantages
Refractive	Wide field of vision; high contrast and high definition	The chromatic aberration is serious; the aperture has a great influence on the performance; the system is bulky and bloated
Reflective	Multi-focus, different numerical aperture and field-of-view angle; spherical and chromatic aberrations; wide spectrum range; simple optical structure, excellent image quality	The production cost is high, the production process is complex, and the off-axis light differs.
Reentrant	Eliminates the refractive-chromatic, reflective-spherical, and other aberrations. The focal length is longer at the same volume	The chromatic aberration is serious, the pore diameter has a great influence on the performance, and the volume is huge. The production cost is high, the production process is complex, and the off-axis light also differs

Cassegrain Optical Antennas

The Cassegrain optical system includes the following optical antennas: large-aperture, array, free-form surface, antenna based on beam-shaping, off-axis dual-reflection, array Gaussian beam off-axis emitting, and pre-collimated. These structures are the results of studying Cassegrain antennas in free-space optical system over the past decades. This study introduces these structures and compares their advantages and disadvantages.

Large-Aperture Cassegrain Antenna

In 1998, AstroTerra [5] designed and developed an optical transceiver antenna for optical communication with the support of the United States Ballistic Missile Defense Organization (BMDO). The antenna was based on a Cassegrain antenna structure, with a weight of 31.5 pounds, volume of less than 1 cubic foot, and a 100-mm aperture. It aligned the transmitting beam, tracked and aligned the signal light, and realized two-way high-speed satellite–ground communications at a maximum transmission rate of 1 Gbit/s.

From 1999 to 2007, the Cassegrain system was used as the main optical antenna in Japan's orbital inter-satellite laser communication satellite [6] (OICETS), the European Aviation Administration's laser inter-satellite link experiment (SILEX) [7], and the United States' optical communication and sensor demonstration (OCD) [8], with antenna diameters of 260 mm, 250 mm, and 100 mm, respectively. The working wavelengths were 1062 nm, 510 nm, and 830 nm, respectively. The maximum powers were 130 W, 115 W, and 90.4 W, respectively.

In 2001, the National Space Development Agency of Japan developed a free-space laser-communication terminal for inter-orbiting satellites [9]. Its uplink working wavelength was 1552 nm, the downlink working wavelength was 1562 nm, the transmitting-antenna aperture was 150 mm, and the maximum speed was 2.488 Gbit/s. When the reception condition was -71.4 dBm, the bit-error rate (BER) was 10^{-6}, the reception sensitivity was 70 photons/bit, and the maximum communication distance reached 25×10^5 km.

In 2013, NASA developed a moon–earth optical wireless communication platform (LLCD) [10]. Its communication terminal on the Moon used a Cassegrain optical antenna with a 108.5-mm optical aperture and a 1550-nm working wavelength. The communication distance of the system reached 45×10^5 km and achieved a two-way communication rate of 20 Mbit/s for the uplink and 615 Mbit/s for the downlink over this distance.

All the above satellite–ground laser communication projects used the Cassegrain optical antenna as the optical system of the transceiver terminal, with no chromatic aberrations and a wide spectral range. Owing to the continuous increase in communication distance, capacity, and speed, the aperture of the Cassegrain optical antenna continues to expand, and the focal length continues to grow. This overcomes the

problem of low received optical-signal energy caused by the middle occlusion ratio of the Cassegrain antenna system and helps meet the actual demand for higher-power optical-signal reception.

Cassegrain Optical-Antenna Array

In 2004, Pitman [11] designed a multi-aperture distributed optical-antenna array, as shown in Fig. 10.4. The system consisted of nine basic array elements with the same focal length. The basic array element was a Cassegrain optical system with a 35-cm aperture. After the nine telescopes were arranged according to a certain layout, they formed a remote-sensing optical-antenna array with a 1.5-m synthetic aperture and 0.8-m focal length. The entire remote-sensing system was composed of a front telescope array, relay optical system, rear support, and servo-control system.

As a subunit, the system could form a larger antenna array with 54 basic arrays. For a 5000-km beam-transmission distance, the maximum resolution of the antenna was 4 m. In addition, it could provide accurate servo control in different bands, distances, and remote-sensing tasks to meet the requirements of accurate imaging. The aperture antenna-array system increases the effective receiving area and the receiving field of view through the expanded synthetic aperture. Moreover, directional reception and interferometry technology make it possible to meet a variety of practical needs.

Freeform-Surface Cassegrain Antenna

In 2008, Takahashi et al. [12] proposed a free-form fiber-coupled Cassegrain antenna with an aiming tracker. A fiber-coupling device made of a glass buckle and optical fiber seamlessly coupled the beam to a single-mode optical fiber. The signal wavelength was 1562 nm, beacon wavelength was 980 nm, and antenna aperture was 40 mm. The beam inclination angle was less than 20 μrad. A tracking collimator with

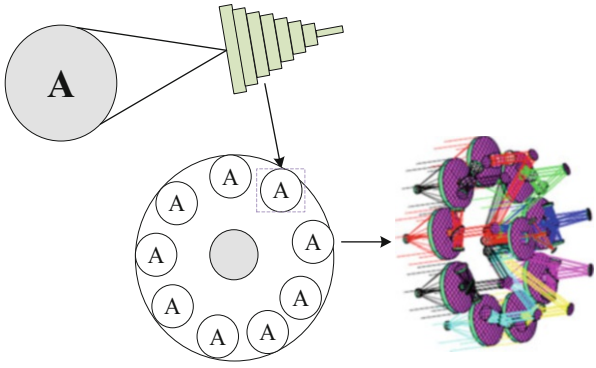


Fig. 10.4 Schematic diagram of a multi-aperture optical-antenna array structure

a response frequency of up to 2 kHz effectively reduced the impact of the angle-of-arrival fluctuation.

In 2011, Li et al. [13] designed a new type of optical antenna for ultraviolet communication. The system structure introduced a double reflector based on a composite paraboloid, which filtered visible and infrared light and reduced the background noise.

In 2019, Gu et al. [14] proposed a design scheme for an unfocused space-laser optical-communication system, based on off-axis free-form surfaces. This scheme adopted a defocused structure to solve the problem of high-power density at the focus of traditional optical-focusing elements when high-power signal sources were used. The simulation results showed that the stellar ratio of the signal beam was greater than 0.8, and the optical field of view was increased by 26.7%, compared to that of a traditional antenna.

Cassegrain Optical Antenna Based on a Beam-Shaping Scheme

In 2002, Kong et al. [15] proposed a design scheme for a ring-beam transmitting optical antenna, as shown in Fig. 10.5. The scheme used two conical or complex-parabolic mirrors to transmit the ring beam, which reduced the beam-energy loss caused by the central block of the card antenna. It reduced the energy loss by up to 45.7%. Owing to the geometric optics and processing level, the accuracy of the prism angle is difficult to guarantee; therefore, the accuracy of this scheme is low, and the error is large.

In 2009, Shishlov et al. [16] designed a Cassegrain optical-antenna structure that could control the beam waveform in a mobile-satellite communication band. As shown in Fig. 10.6, the antenna is composed of two mirrors and a phased-array controller. The main reflector is used to improve the antenna gain, and the secondary reflector rotates in elevation and azimuth to control the beam. The phased array maximizes the radiation direction (Fig. 10.7).

In 2012, Yu et al. [17] proposed a scheme that converted a solid Gaussian beam into a hollow Gaussian beam by correcting a pair of two-dimensional optical-

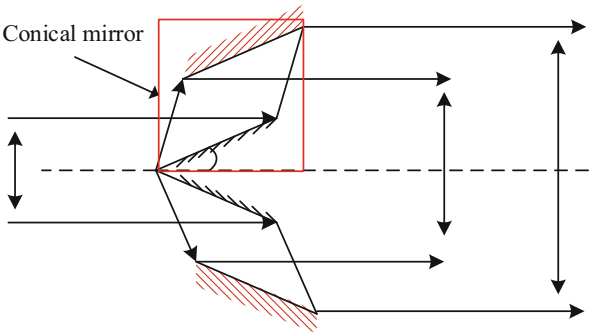


Fig. 10.5 Schematic diagram of a conical-mirror structure

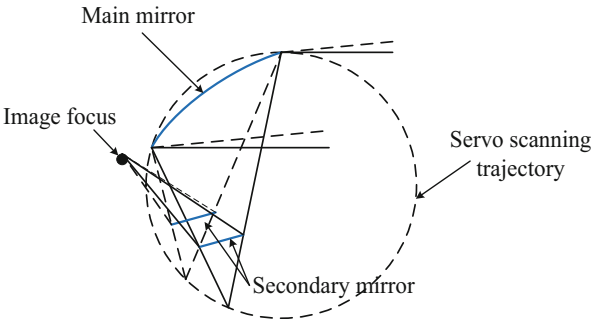


Fig. 10.6 Schematic diagram of a waveform control-card antenna

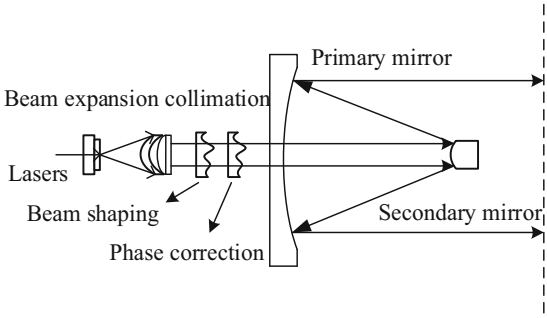


Fig. 10.7 Structural diagram of an annular-beam emission optical antenna

diffraction elements through beam shaping and phase compensation. This avoided the beam-energy loss caused by central occlusion. The emission efficiency of this optical-system structure exceeded 95%.

In 2018, Xu et al. [18] designed a new type of Fresnel optical antenna. As shown in Fig. 10.8, the signal beam is first incident to the phase-modulated Fresnel zone plate at the center of the main mirror, and the outgoing beam is a hollow beam with a focus. The beam is incident on the secondary mirror in the hyperbolic structure; it is then reflected by the primary mirror and emitted parallel to the antenna structure. The system uses a self-designed phase-modulated Fresnel zone plate.

The transmission efficiency of the Gaussian beam, in the secondary-mirror occlusion part at the exit pupil of the antenna, was 92.82%, with an efficiency increase of 31.42%. Compared to the diffraction-element annular Gaussian-beam emission scheme, the zone plate has a simple structure and is easy to operate and realize. Although the emission efficiency is slightly lower, it has good practicability and realizability.

The above four beam-shaping schemes are proposed to compensate for the low light-energy utilization rate of the Cassegrain antenna. The structure and principle of the scheme proposed in Reference [19] are relatively simple, but the requirements for

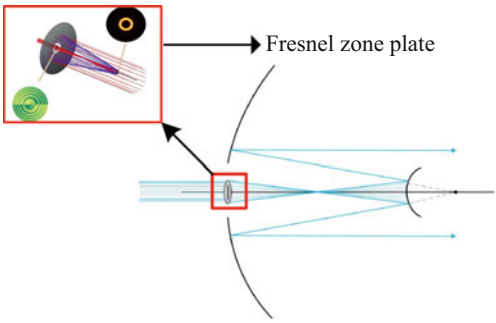


Fig. 10.8 Structural diagram of a ring-beam emitting optical antenna

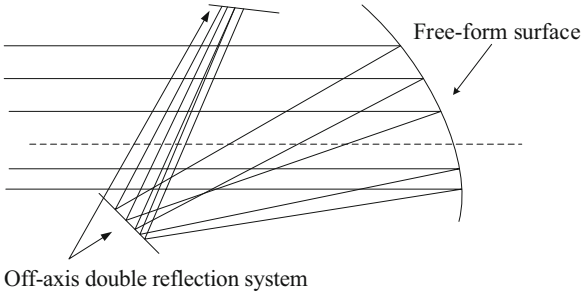


Fig. 10.9 Schematic diagram of an off-axis double-reflection structure with a free-form surface

the transmittance and reflectivity of the prism are very high. It is difficult to meet the manufacturing-process requirements.

The scheme proposed in References [23, 24] uses the phase-shaping principle to output the beam as a circular beam. This is more suitable for practical applications, owing to the good dispersion and temperature-difference characteristics, high diffraction efficiency, high accuracy, and operability of the binary-diffraction element. The scheme proposed in Reference [13] receives optical signals in a larger field of view and adjusts the antenna for different angles and beam conditions to achieve optimal reception. This expands the receiving field of the Cassegrain antenna.

Off-Axis Dual-Reflection Cassegrain Optical Antenna

In 2017, Shu et al. [19] proposed using an off-axis double-mirror reflection structure with a free-form surface instead of a coaxial double-mirror reflection structure for space laser-communication optical antennas. As shown in Fig. 10.9, the light emitted through the free surface accurately entered the first coated mirror of the off-axis double-reflection system, and the reflected light entered the optical-relay system.

In 2019, Tian et al. [20] designed a four-mirror optical antenna with a long exit-pupil distance axis. The exit-pupil position of the optical system coincided with the

position of the fast mirror. This eliminated the optical relay system and overcame the defect that the backscattered light of the optical relay system adversely affected the laser-communication terminal.

According to the above optical antenna structure, we can conclude that the addition of a free-form surface can effectively correct off-axis aberration. Compared to a coaxial transmission system, the modulation transfer function (MTF) of the off-axis reflection system is higher, and it can work closer to the diffraction limit in a large field of view.

With the constant expansion of the field of view, the performance of the coaxial system decreases rapidly, whereas the off-axis dual-reflection structure remains stable and can approach the diffraction limit. In addition, the off-axis dual-reverse structure avoids the problem of low energy utilization caused by the key occlusion of the Cassegrain optical ensemble. This makes it more suitable for spaceborne laser communications and improves the overall performance of the space laser-communication system.

Maksutov–Cassegrain Optical Antenna with Off-Axis Emission of a Gaussian-Beam Array

In 2015, Ke and Wang et al. [21] designed an array of off-axis Gaussian-beam transmitting Maksutov–Cassegrain antennas in response to the optical-signal energy waste caused by the occlusion ratio of the Maksutov–Cassegrain antenna (hereinafter referred to as the Maka antenna). Because the coaxial transmission efficiency of the Maka antenna was low, and the reception efficiency was low because of the secondary-mirror occlusion, the Maka antenna structure was properly modified so that it could emit Gaussian beams off-axis. Moreover, its transmission and reception efficiencies were further improved through the self-coupling phenomenon of the Gaussian-beam array.

The emission scheme for arranging the Gaussian beams proposed is shown in Fig. 10.11. In a rectangular coordinate system, it is assumed that the Gaussian beams are uniformly distributed on a circle with radius r_0, and the angular spacing of the adjacent subbeams is. The subbeam constitutes a radially distributed beam array (Fig. 10.10).

After the derivation and calculation, a Gaussian-beam array coupling-simulation experiment was conducted. According to the experiment and results, if the diverging Gaussian-beam array ($F_0 < 0$), the collimated Gaussian-beam array ($F_0 = \infty$), and the converged Gaussian-beam array ($F_0 > 0$) are horizontally transmitted along the optical axis, the Gaussian beam will always change because of its wavefront curvature as the transmission distance increases. Finally, it evolves from a discrete Gaussian beam into a beam. The farther the distance, the better the beam-coupling effect, and the greater the central-energy attenuation.

Researchers modified the conical aperture of the Maka antenna and used a microlens to control the parameters of the Gaussian beam emitted by a single-mode optical fiber. They designed four Maka antennas with uniformly distributed

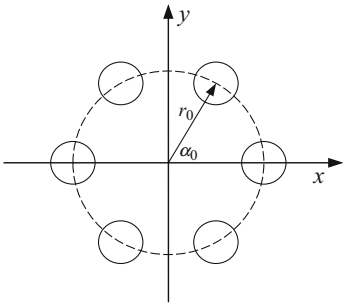


Fig. 10.10 Schematic diagram of a radially distributed beam array [22]

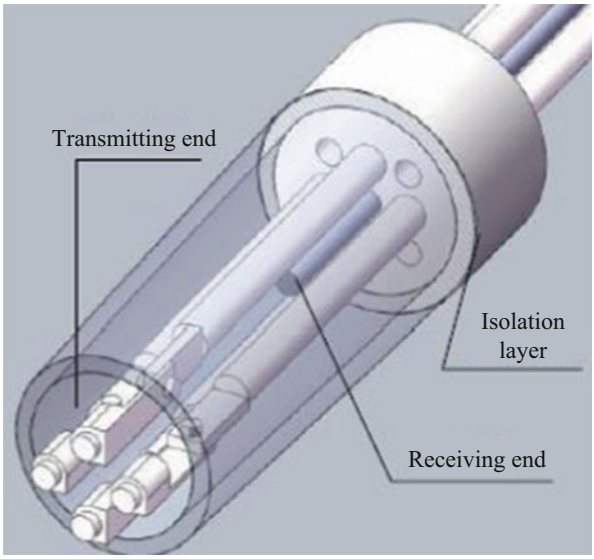


Fig. 10.11 Schematic diagram of an off-axis transmitting antenna with a uniform circular arrangement of Gaussian beams

off-axis emissions of Gaussian beams in a ring. The structure is shown in Fig. 10.11 [21].

Experiments were performed on the coupling effect of the off-axis emitting beam array. Figure 10.12 shows the spot emissions of the beam at different transmission distances [21]. The experimental results are listed in Table 10.2. From the experimental phenomenon in Fig. 10.12 and the experimental data in Table 10.2, it can be seen that the farther the distance is, the larger the light spot is, and it gradually approaches the center. This is because the emitting beam has become a divergent Gaussian beam and gradually self-couples.

To meet the requirements of duplex wireless laser communication, Si et al. [23] made certain structural modifications to the Maka antenna: the optical signal was connected to the semiconductor-laser output by four pigtails through the standard

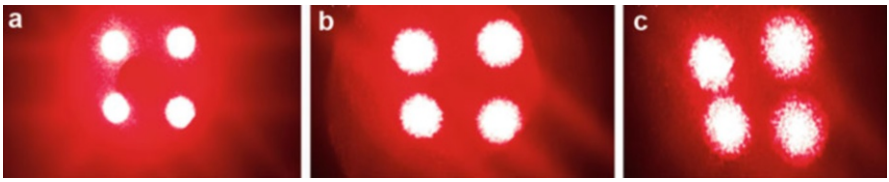
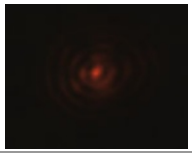
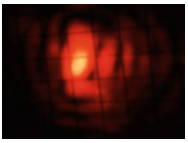
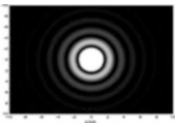


Fig. 10.12 Spot pattern of a Gaussian beam array. (**a**) 10 m (**b**) 50 m (**c**) 100 m

Table 10.2 Experimental results of the off-axis emission of a single Gaussian beam

	Normal launch (experiment)	Off-axis emission (experiment)	Off-axis emission (theory)
Exit pupil	9%	53%	100%
Efficiency	0	2.2°	$0.25 \times 10^{-3\circ}$
Light-spot shape			

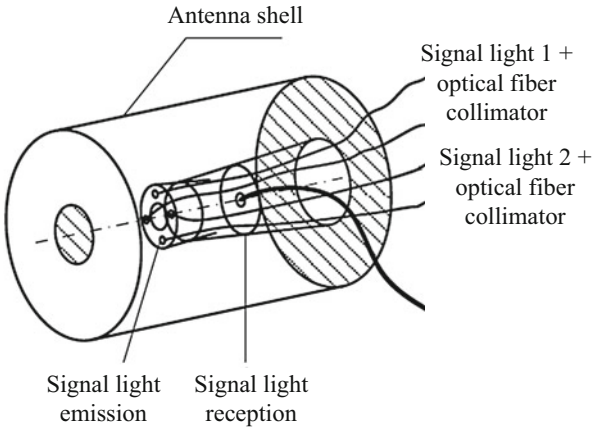


Fig. 10.13 Schematic diagram of the integrated transceiver system

interface. A fiber collimator was installed at the rear end of the pigtail to change the divergence angle of the pigtail beam. According to the space position of the transmitting optical fiber, the four pigtails were fixed on the circular transmitting optical-fiber disk to complete the transmitting function, as shown in Fig. 10.13.

In addition, the optical fiber was placed in the center of the receiving optical disk. The received optical signal was coupled into the optical fiber in the center of the receiving optical disk after passing through the through hole of the transmitting optical disk. The receiving optical fiber disk was placed at the focal point of the

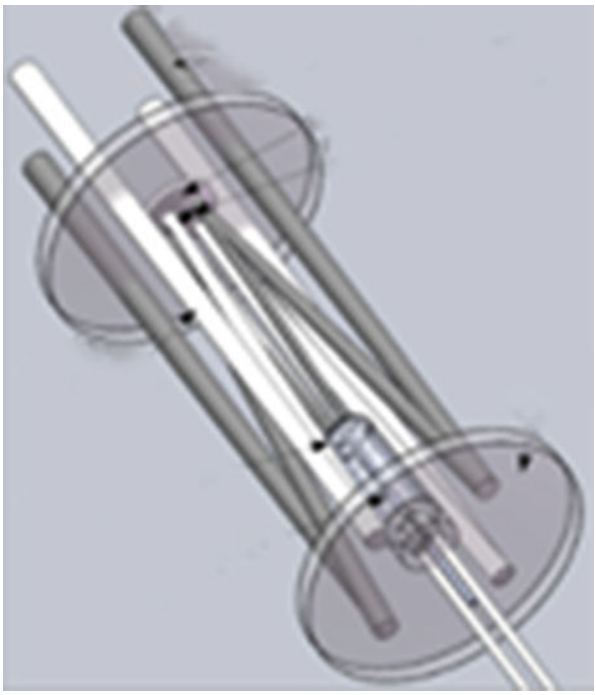


Fig. 10.14 Three-dimensional effect of sending and receiving

Table 10.3 Experimental data of Maka antennas A and B

Distance (m)	Emission optical power (B)	Receive optical power (A)	Coupling efficiency (A)	Emission optical power (A)	Receive optical power (B)	Coupling efficiency (B)
5	16.8 mW	3.1 mW	21.6%	16.8 mW	2.8 mW	20.2%
15		3.0 mW	20.9%		2.9 mW	20.9%
30		2.3 mW	16.1%		2.2 mW	18.8%

Maka antenna in the conical aperture to complete the spatial optical-fiber coupling and optical reception. The three-digit effect is illustrated in Fig. 10.14.

The experimental results of the off-axis emission tests of the Gaussian beam are shown in Table 10.3. Maka antennas A and B were used as transmit and receive antennas, respectively, to test the off-axis transmission optical power and the reception and coupling efficiency. The experimental results showed that the transmitted optical power was 16.8 mW after passing through the collimator. The emitted optical power was 14.9 mW after passing through Maka antenna A, and the transmission efficiency was up to 94%.

Compared with the traditional optical-antenna coaxial-transmission method, the transmission efficiency was significantly improved. This optical antenna structure can solve the optical-power loss problem caused by secondary-mirror occlusion.

Pre-Collimated Cassegrain Optical Antenna

In 2006, Yang et al. [24] designed a new type of aspheric pre-collimated Cassegrain optical antenna for a space laser-communication system. The system aperture was 150 mm and the focal length of the aspheric lens was 250 cm. The system structure compressed the beam to a minimum of six μrad with a 20× magnification, but it also improved the light-emission gain and reduced the background-light noise.

Jiang et al. [25] proposed to use optical devices, such as collimation systems, optical isolators, and splitter prisms, to integrate transceivers after focusing the Cassegrain optical antenna. This is the most common method of transceiver integration.

Further improvement and research on the traditional Cassegrain optical antenna is mainly intended to reduce the impact of the middle obstruction of the Cassegrain antenna structure. Various technologies and structures have been proposed, such as ring-beam emission, hollow-beam emission, off-axis emission of beam arrays, diversity reception, and coaxial aperture synthesis. These technologies have advantages and disadvantages in terms of their structure and function, as shown in Table 10.4.

In recent years, many new optical antennas have emerged, such as the complex parabolic-reflector optical antenna, the combined Galileo and Gregorian optical antennas, complex aspheric-lens combination, and bionic compound-eye structure optical antenna. These new structures inherit the performance of traditional card antennas, but have advantages, such as high gain and aberration elimination.

However, domestic research works mainly focus on numerical simulations of the antenna structure and design schemes [26], lacking practical engineering applications. Only a few researchers have conducted test and application cases, such as the Institute of Optoelectronic Technology of the Chinese Academy of Sciences [27] and Xi'an University of Technology [28].

Beacon-Free Tracking

The signal beam in an optical wireless communication system usually has a very small beam-divergence angle to ensure sufficient detection-signal strength. However, the small divergence angle is not conducive to the acquisition, pointing, and tracking (APT) of the receiver. Therefore, a beacon light with a large beam-divergence angle and a signal light with a small divergence angle are sent along the optical path to facilitate APT at the receiving end. A beacon light requires an independent optical-transmission system, which complicates the optical structure and servo control system of the transceiver; therefore, no independent-beacon optical APT systems exist.

An optical APT system without an independent beacon divides part of the signal light into a spot-position detector to detect the beam position [29]. The terminal-signal and beacon beams of the non-independent beacon optical APT system share

Table 10.4 Research progress on the Cassegrain antenna structure

Antenna structure	Research progress	Advantages and disadvantages
Large aperture	Large-caliber technology	The large-aperture antenna can increase the power of the received optical signal and uses the aperture-average effect to weaken the fluctuation
Array	Large-caliber technology	It breaks through the maximum-aperture limit of a single antenna. The array gain can be obtained through signal coherence. Multiple basic-unit channels can obtain diversity gain and eliminate inter-channel interference
Free-form surface	Free-form surface-mirror technology	The energy loss of the coaxial-reflection antenna is reduced and the transmission efficiency is improved
Antenna structure based on beam shaping	Beam-shaping technology	The divergence-angle collimation shaping can minimize the light spot at the receiving end. The light-field energy-distribution shaping reduces the inherent energy loss of the antenna.
Antenna structure based on beamforming	Off-axis reflection technique	The light-collecting ability of the antenna is improved and its detection range is increased. Considering the technology maturity, system performance, and future communication requirements, the off-axis double-inverse system has higher feasibility and maneuverability
Off-axis emission structure of a Gaussian-beam array	Beam auto-coupling and technology	It effectively solves the transmission-efficiency decline caused by the secondary mirror reflection of the Maka antenna
Pre-collimating	Lens pre-collimation technology	The collimating lens further compresses the divergence angle of the beam, improves the spot quality, and improves the receiving and transmitting efficiency

parts of the optical elements. Therefore, this type of receiving antenna has a small size, light weight, and low power consumption.

In 2007, the Tesat-Spacecom communication satellite launched by the German Aerospace Technology Center used the then-highest level of beaconless capture technology [30]. In 2017, Guo et al. [31] proposed a beacon-free tracking and acquisition method that required 0.135 W of laser emission power, had 30 s of acquisition time, and had a 95% acquisition probability.

In 2018, a satellite launched by the European Space Agency was equipped with the OPTEL-D communication terminal with a 1016-mm receiving aperture and a 1550-nm receiving wavelength. The terminal used a reflective-telescope structure combined with a multi-axis beaconless optical APT tracking system [32]. The beacon-free tracking and aiming system avoids the shortcomings of a beacon optical APT system, does not require an additional beacon optical

antenna, greatly simplifies the complexity of the transmitter and receiver, and is the development trend of future APT systems.

Point-to-Multipoint Communication Optical Antennas

The above antenna structures are suitable for point-to-point optical wireless communication. However, the limited number of communication terminals in point-to-point communication systems cannot carry out high-speed communications and information relays between multiple platforms. This restricts the progress of space laser-communication applications and the development of laser-communication space networking. Therefore, the networking mode of point-to-point laser communication has become a popular research topic.

In 1993, Clarke et al. [33] proposed to use a large field-of-view optical antenna to fully cover low-orbit users and installed a movable micro detector on the antenna focal plane for point-to-multipoint laser communication.

In 2002, the American Transformational Satellite Communications System (TSAT) proposed placing multiple transceivers in different locations to achieve simultaneous multipoint communication [34].

In 2005, Sherman et al. [35] proposed placing an optical-fiber array at the focal plane of a receiving antenna to realize point-to-multipoint space laser communication.

Although these schemes designed one-to-many communication links, the communication quality is not ideal, owing to the limited effect of the optical-fiber numerical aperture on the received signal. In addition, the arrangement of the optical fiber and optical-fiber array has some difficulties. Therefore, the practicability and integrated miniaturization of one-to-many space optical communication antennas must be further explored.

In 2006, Nishi et al. [36] developed a 60-GHz point-to-multipoint wireless-access link with a 156-Mbit/s data rate for short-range applications, such as indoor optical wireless local area networks and intelligent-transportation systems. The total size of the 60-GHz transceiver module was 50 mm $\times$ 75 mm.

In 2011, Jiang et al. [37] proposed a point-to-multipoint optical antenna with a paraboloid-based multimirror splicing structure. The antenna structure is composed of front, middle, and rear lenses whose surfaces are concentrated. The communication target within the 160° field-of-view angle can be imaged on the bottom receiving-end plane, and the signal beam can be transmitted to communicate with other targets in this range. The communication rate of the antenna can reach 2.5 Gbit/s and the pitch angle is $\pm 60°$.

In 2014, Zheng et al. [38] proposed an optical-antenna structure with four mirrors on the orbital surface and proposed a link satellite-networking scheme based on this antenna structure. The antenna structure captures the laser light source within a 50° view angle and 40° pitch angle. The servo system enables it to conduct two-way communication and dynamic tracking between satellites in the same and different orbits. The pitch angles of the antenna devices in Beijing,

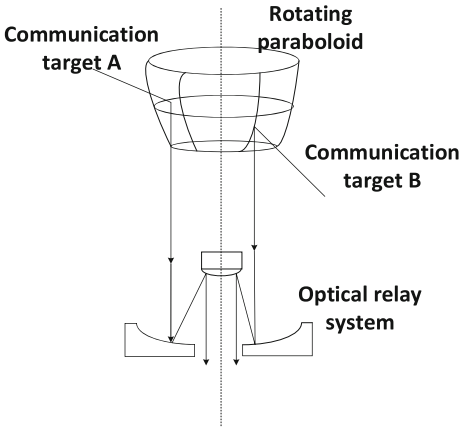


Fig. 10.15 Structural diagram of an optical antenna for rotating paraboloid one-to-many laser communication

Mohe, Lhasa, and Shanghai met the requirements of a 1500-km minimum communication distance.

In 2015, Jiang et al. [39] proposed an optical-antenna design scheme that met the requirements of synchronous duplex laser communication between multiple points, to enable networking communication between multiple platforms in orbit. The structural diagram is shown in Fig. 10.15. The optical antenna consists of a rotating paraboloid mirror and an optical relay system. The aperture of the rotating paraboloid mirror is 200 mm and the pitch angle is 73°.

Communication targets A and B are in different directions. The light is incident onto the optical system along the focal direction of the parabolic surface and along the direction parallel to the principal axis. It then enters the optical relay system.

The antenna has the advantages of a large field of view, wide communication range, simple structure, and high receiving energy. It can realize simultaneous duplex communication between multiple points and can meet the requirements for networking communication between geostationary-orbit satellites (GEO) and their optical ground stations (GEO-OGS) [40].

In this study, Jiang et al. also proposed a point-to-multipoint laser-communication optical tracking scheme that used a rotating paraboloid and a three-concentric-sphere structure as the receiving antenna. It simultaneously communicated with up to three systems in the experiment.

A schematic diagram of the three-concentric-sphere structure is shown in Fig. 10.16. The structure has a 120° field-of-view angle in both the horizontal and pitching directions. The optical antenna can perform simultaneous duplex communication with multiple transceiver terminals on the track of the corresponding working area. When the transceiver terminal moves along the track, out of the working area of the current optical antenna and into the next area, the next optical antenna takes over the transceiver terminal and establishes a communication link.

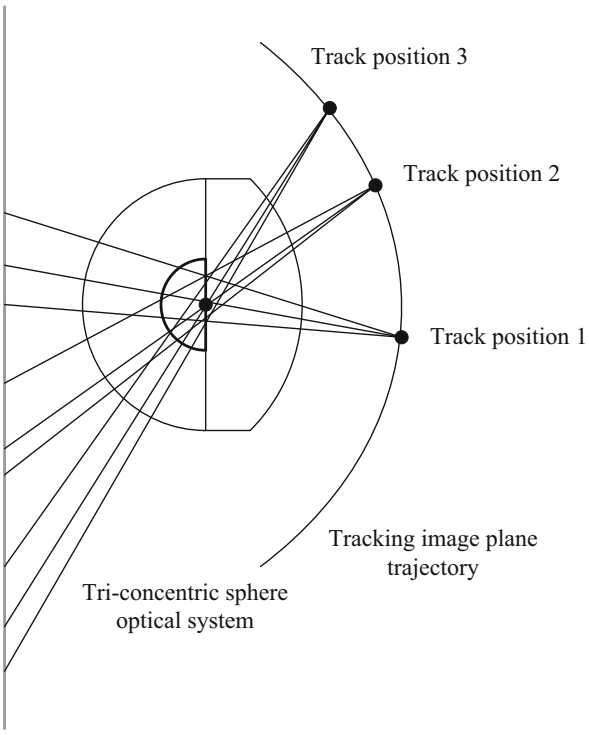


Fig. 10.16 Schematic diagram of a concentric-sphere structure

This scheme omits the complex steps of tracking and aiming and provides a feasible scheme for establishing a multipoint simultaneous duplex tracking communication network.

In 2015, Zhang et al. [41] designed an optical antenna for an optical wireless communication network that could realize synchronous duplex space-laser communication and carry out bi-directional communication between multiple points in a 360° azimuth angle and 90° pitch angle range. The highest transmission rate was 2 Gbit/s.

In 2016, Jiang et al. [42] proposed using a wide-angle beam expander and a double-wedge group to make the optical antenna of a multipoint laser communication system. As shown in Fig. 10.17, a 350-mm wide-angle beam expander expands the field of view and receives optical signals from different directions. The double optical-wedge group with 100-mm *N*-pair apertures can dynamically aim and track *N* different signal sources simultaneously and can carry out one to three simultaneous two-way communications in a low track.

In the communication experiment, the communication distance was 400 km, communication rate was 2.5 Gbit/s, and communication bit-error rate was less than 10^{-7}. The optical antenna flexibly controlled the relative position of the number

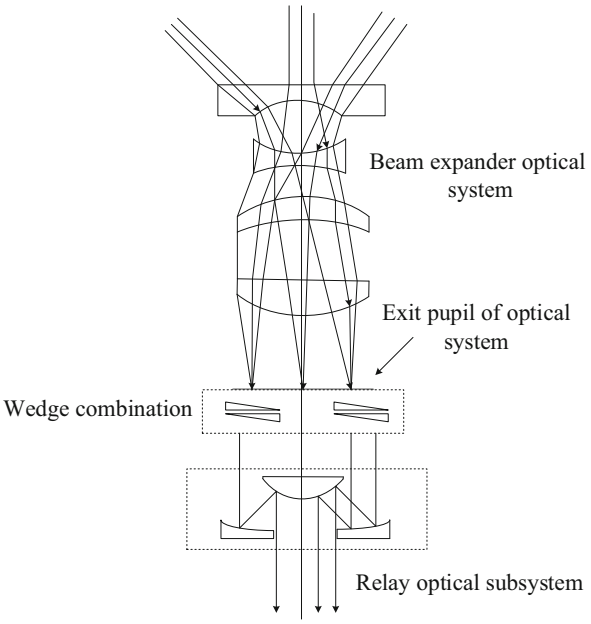


Fig. 10.17 Structural diagram of a combined wide-angle beam expander and double optical wedge

of optical-wedge combinations, according to the number and position of communication targets, and used as much optical-signal energy as possible.

Compared with the three-concentric-sphere structure, the structure of the combined double optical-wedge and rotating-paraboloid optical antenna is slightly complex, and its weight and volume are higher. However, its optical-energy utilization rate is higher, making it more suitable for long-distance optical wireless communication-system networking.

A comparison of the point-to-multipoint antenna structures is presented in Table 10.5.

Establishing a space-to-ground integrated global information network is inevitable. Traditional point-to-point space laser communication cannot communicate simultaneously with multiple terminals, which limits the networking ability of optical wireless communication networks. Therefore, point-to-multipoint optical wireless communication is the only way to develop networking technology.

Summary

This chapter systematically summarized the application and development of optical-antenna technology in optical wireless and space-laser communication systems. In the future, the transmitting/receiving antennas of optical wireless communication systems still need to be developed.

Table 10.5 Progress of optical antennas for point-to-multipoint laser communication

Year	Researcher	Research progress	Advantages and disadvantages
1993	Clarke	Full-coverage optical antenna with a large field of view	(1) The field of view is still small, owing to the numerical-aperture limitation of the optical-fiber and card system (2) The stray light is serious and the device arrangement and combination are difficult, with low application and feasibility. Further research is needed
2002	TSAT	Multi-receiver combined with an optical communication antenna	
2005	Sherman	Focal-plane fiber array with multiple receiving antennas	
2006	Nishi	Indoor optical wireless point-to-multipoint wireless-access optical antenna	
2011	Jiang H	Splicing paraboloid optical antennas with multiple mirrors	(3) Wide communication range, simple structure, high usable aperture, and high beam energy (4) The communication range is long, so it is suitable for multi-terminal networking of aircraft, satellites, and ground terminals
2014	Zheng	Optical antenna with four mirrors on track surface	
2016	Jiang L	Wide-angle beam expander and double-wedge multi-point receiving optical antenna	
2015	Jiang H	Multi-point receiving optical antenna with three-concentric-sphere structure	(5) The concentric sphere has a simple structure, light weight, and miniaturization, which is suitable for short-range and flexible laser communication networks
2015	Zhang	Design scheme of an optical antenna for optical wireless communication network	

1. The theory and technology of point-to-multipoint optical antennas should be investigated to enable optical wireless communication networks.
2. The traditional APT beacon requires an independent optical antenna for transmission, which complicates its structure. The beacon-free optical scheme saves a light source, which simplifies the transmitter structure, and can meet the miniaturization and integration requirements of future optical devices.
3. Multi-emitting and multi-receiver optical antennas should be investigated to further improve the flexibility of optical wireless communication networks. This can enable a space-to-ground three-dimensional laser-communication network and expedite the practical application of wireless laser communications.

References

1. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress of adaptive optics in Optical Wireless Communication system for Xi'an University of Technology [J]. High Power Laser and Particle Beams, 2021, 33: 081003.
2. Sun Wenchang, Yuan Dan. Research and Development of wide viewing Angle Optical receiving Antenna [J]. Volkswagen Technology, 2012, 14(6):54–58.
3. Zhang Yuxia, Ai Yong. Discussion on the disadvantages of Cassegrain antenna based on space optical communication [J]. Infrared and Laser Engineering, 2005, (5):60–63.
4. Li Xiang, Lan Tian, Wang Yun, Wang Long-Hui. Design and study of Fresnel lens for an antenna in indoor visible light communication system. Acta Phys. Sin., 2015, 64(2): 024201.

5. Korevaar E J, Hofmeister R J, Schuster J J, et al. Design of satellite terminal for Ballistic Missile Defense Organization (BMDO) lasercom technology demonstration [C]//Free-space laser communication technologies VII. SPIE, 1995, 2381: 60–71.
6. Yuuichi Fujiwara, Masaaki Mokuno, Takashi Ono, et al. Optical inter-orbit communications engineering test Satellite (OICETS) [J].ActaAstronautica, 2007, 61(1):163–175.
7. Tolkert, Nielson B, Demelenne E. Desplats In-orbit test results of the first SILEX terminal [J]. SPIE Free-Space Laser Communication Technologies, 1999, 36(15):31–42.
8. Tsunyee Y, Muthu J, James R. Progress on the development of the optical communications demonstrator [J]. Photonics West, 1980, 123(10):94–101.
9. Toyoshima M, Yamakawa S, Yamawaki T, et al. Reconfirmation of the optical performances of the laser communications terminal onboard the OICETS satellite [J]. Acta Astronautica, 2004, 55(3–9): 261–2610.
10. Song Tingting, Ma Jing, Tan Liying, Yu Siyuan, Ran Qiwen. Lunar Laser Communication Demonstration in USA: Terminal Design [J]. Laser & Optoelectronics Progress, 2014, 51(5): 050003.
11. Pitman J T, Duncan A, Stubbs D, et al. Remote sensing space science enabled by the multiple instrument distributed aperture sensor (MIDAS) concept [C]. Instruments, Methods, and Missions or Astrobiology VIII. International Society for Optics and Photonics, 2004, 5555: 301–311.
12. Takahashi K, Arimoto Y. Compact optical antennas using free-form surface optics for ultrahigh-speed laser communication systems [J]. Optical Engineering, 2008, 47(1):16–22.
13. Li Xiaoming, Hua Wenshen, Wu Xianquan. Research on optical antenna for ultraviolet communication based on double reflector [J]. Optical Instrument, 2012, 34 (2): 21–25.
14. Gu Sisi, Cui Zhangang, Qi Bo. Design of optical antenna for laser communication based on off-axis free-form surface [J]. China Optics, 2020, 13 (3): 547–558.
15. Kong Xianglei, Hao Peiming. A new reflective laser beam expansion scheme to eliminate central obstruction [J]. Journal of Quantum Electronics, 2002, 21 (3): 205–2010.
16. Jung Y B, Shishlov A V, Park S O. Cassegrain antenna with hybrid beam steering scheme for mobile satellite communications [J]. IEEE Transactions on Antennas and Propagation, 2009, 57(5): 1367–1372.
17. Yu Jianjie, Han Qiqi, Ma Jing, Tan Liying. Potential application of diffractive optical elements in satellite laser communication terminals [J]. Infrared and Laser Engineering, 2013, 42(1): 130–138.
18. Xu Fangya. Design of new Fresnel optical antenna [D]. Chengdu: University of Electronic Science and Technology, 2019.
19. Shu G, Chao W, Shoufeng T, et al. Optical antenna design of off-axis two-mirror reflective telescope with freeform surface for space laser communication [J]. Infrared and Laser Engineering, 2017,46 (12): 1222003–1222003 (8).
20. Qijie Tian, Jianhua Zhang, Huanhuan Zhang, Yansong Song, Jingjing Xue, Xinrong Hu, Xiang Chen, Shuai Li. Design of Off-Axis Four-Mirror Optical Antenna with Long Exit Pupil Distance [J]. Acta Optica Sinica, 2020, 40(18): 1822001(in chinese).
21. Ke Xizheng, Lei Sichen, Wang Tianyu. Cassegrain transceiver integrated antenna based on fiber array [J]. Progress in Laser and Optoelectronics, 2015, 52 (12): 49–55.
22. Lei Sichen, Ke Xizheng, Shao Junhu, et al. Experimental study on automatic alignment of space optical-fiber array coupling [J]. Laser Technology, 2014, 38 (2): 191–195.
23. Sic L, Xiz K. Laser emitted through the off-axis Cassegrain antenna [C]. International Symposium on Photoelectronic Detection and Imaging 2013: Laser Communication Technologies and Systems, 2014.
24. Yang H J, Hu Y, Li C, et al. Optimum design for optical antenna of space laser communication system [C]. 2006 International Conference on Communications, Circuits and Systems. IEEE, 2006, 3: 2016–20110.
25. Jiang Ping. Design and simulation of active and passive multi-wavelength optical antenna [D]. Chengdu: University of Electronic Science and Technology, 2008.

26. Jiang Wenhan, Zhang Yudong, Rao Changhui, et al. Research progress of adaptive optics in the Institute of Optoelectronic Technology, Chinese Academy of Sciences [J]. Journal of Optics, 2011, 31 (2): 0900106.
27. Yang Huizhen, Li Xinyang, Jiang Wenhan. Application progress of adaptive optics technology in atmospheric optical communication system [J]. Progress in Laser and Optoelectronics, 2007, 44 (1): 61–69.
28. Ke Xizheng, Wu Jiali. Principle and application of wireless optical coherent communication [M]. Beijing: Science Press.
29. Li Qing. Research on weak beacon optical signal processing technology of ATP system in space optical communication [D]. Chengdu: University of Electronic Science and Technology, 2020.
30. Whiteley M R, Goorskey D. Influence of aero-optical disturbances on acquisition, tracking, and pointing performance characteristics in laser systems [J]. Proceedings of SPIE, The International Society for Optical Engineering, 2011, 8052(3): 298–306.
31. Guo Aiyan, Gao Wenjun, Zhou Aosong, etc Beacon free acquisition and pointing technology for inter-satellite optical communication [J]. Infrared and Laser Engineering, 2017, 46 (10): 7–11.
32. Stumpf M, Roth C, Mosberger M, et al. OPTEL-D: an optical communication system for the deep space [C]//Free-Space Laser Communication and Atmospheric Propagation XXX. International Society for Optics and Photonics, 2018, 10524: 105240D.
33. Clarkee S, Prenger R, Ross M. Experimental results with a prototype three-channel multi-access transceiver lasercom terminal [C]. SPIE, 1993, 1866:128–138.
34. Zhizhong Y, Jingyi C, Xianwei Z. Development and technology of U.S. military satellite communications systems [J]. Communication Technology, 2009, 41(11): 55–59.
35. Sherman M P, John A T. Point-to-Multipoint Free-Space Optical Wireless Communication System [P]. US Patents,6445496B1, 2002-9-3.
36. Choi S T, Yang K S, Nishi S, et al. A 60-GHz point-to-multipoint millimeter-wave fiber-radio communication system [J]. IEEE transactions on microwave theory and techniques, 2006, 54(5): 1953–1960.
37. Jiang Huilin, Hu Yuan, Song Yansong, et al. Research on optical transceiver technology for space laser communication networking [J]. Space Return and Remote Sensing, 2011, 32 (5): 52–510.
38. Zheng Yang, Fu Yuegang, Hu Yuan, et al. Research on dynamic alignment of four mirrors based on space laser communication networking [J]. China Laser, 2014, 41 (1): 103–1010.
39. Mojtaba M A, Zabih G, Stanislav Z, et al. Dual-purpose antenna for hybrid free space optics/RF communication systems [J]. Journal of Lightwave Technology, 2016 (Web publishing).
40. Zhang Tao, Fu Qiang, Li Yahong, et al. Design of multi-platform laser communication networking confocal reflective optical antenna [J]. Journal of Optics, 2015, 35 (10): 220–226.
41. Purvinskis R, Giggenbach D, Henniger H, et al. Multiple-wavelength free-space laser communications [C]//SPIE. Free-Space Laser Communication Technologies XV. SPIE, 2003, 4975: 12–110.
42. Jiang Lun, Hu Yuan, Wang Chao, et al. Research on optical system of point-to-multipoint simultaneous space laser communication [J]. Journal of Optics, 2016, 36 (5): 37–43.

Research Progress on One-to-Many Transmitting Antennas for Optical Wireless Communication

11

Contents

Introduction 322
Optical Antennas 323
Reflective Telescope Systems 323
Refractive Telescope Systems 326
Catadioptric Telescope System 328
Problems of Point-to-Point Antennas 330
Foreign and Domestic Research Progress 330
One-to-Many Optical Communication 336
Basic Principles of One-to-Many Optical Communication 336
Acquisition, Pointing, and Tracking (APT) 340
Optical Principle of One-to-Many Communication 341
Multi-optical Terminal Combination Type 342
Multi-reflector Type 342
TSAT Networking 343
Double Optical-Wedge Group Combined Type 344
Fiber-Array Type 345
Multi-transmission, Multi-receiver Structure 345
Three-Concentric-Sphere Structure 347
Rotating-Paraboloid Structure 348
Development Trends 348
Summary 350
References 351

Abstract

Optical antennas play the role of transmitting and receiving coupled optical signals in optical wireless communication, and a pair of multipoint transmitting antennas can expand optical wireless communication applications. This chapter summarizes the development of one-to-multipoint optical antennas at home and abroad and introduces the basic principles of optical fiber array structure, three-concentric sphere structure, and rotating paraboloid one-to-multipoint optical antenna. Finally, the development prospects of this field are foreseen. Optical

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_11

wireless communication is an emerging communication technology with advantages of high speed, large capacity, and good security. However, the traditional optical wireless communication system has the limitation of single-to-single communication mode, which can't satisfy the demand of multi-user simultaneous communication. Optical antenna plays the role of transmitting and receiving coupled optical signals in optical wireless communication, and a pair of multipoint transmitting antennas can expand the applications of optical wireless communication. The development of one-to-multipoint optical antennas at home and abroad is summarized, and the basic principle of one-to-multipoint optical antennas is introduced. Several existing one-to-multipoint optical antennas design schemes are discussed. Finally, the development trend of this field and the future development direction are prospected.

Keywords

Optical wireless communication · Optical antennas · One-point to multi-point · Basic principles

Introduction

An optical antenna is a telescopic system that can be divided into three categories, depending on the angle of incidence of the optical antenna:

1. Reflective telescopic systems, which reflect incident light
2. Refractive telescopic systems, which refract light
3. Reflexive telescopic systems, which refract and reflect light [1]

Transmitting antennas concentrate the energy from a light by gathering it into a small angular range and transmitting it into the atmosphere to extend the communication distance. Receiving antennas, on the other hand, use various optical elements to increase their receiving area. They improve the signal-to-noise ratio by receiving as much light signal and energy as possible to reliably receive the signal. The amount of received energy has a direct impact on the bit-error rate (BER) of the system [2].

The selection of transmitting and receiving antennas is very important and directly affects the communication quality of the entire system. In practical applications, multiple transmit and receive antennas are typically used to improve the communication quality and stability for multi-channel transmission and reception.

Point-to-point communication systems have a limited number of communication terminals, which greatly limits the interstellar and space-to-ground transmission of high-capacity and high-rate information. This seriously affects communication relays, network formations, and applications. Therefore, the development of point-to-multipoint communication has become an inevitable trend in the future.

Optical Antennas

Optical antennas are classified as shown in Fig. 11.1. Reflective telescope systems are widely used in astronomy, earth science, and other fields because of their special construction, which can be applied to various wavelength bands, including visible light, infrared, and others.

Refractive telescope systems are typically used for optical wireless communication in the visible and near-infrared wavelengths. Their simple construction and low cost make them widely used in some low-cost short-range communication systems.

The Catadioptric telescope system, on the other hand, is developed by combining the advantages of the above two telescope systems. Owing to its aberration correction, the Catadioptric telescope system is not only suitable for optical wireless communication in visible and near-infrared wavelengths but also performs particularly well in long-range optical wireless communication in complex atmospheric environments [3].

Reflective Telescope Systems

Reflective telescopic systems use mirrors, both spherical and aspherical. They use the principle of specular reflection, rather than lens refraction, to achieve optical imaging, and have multiple focal points, so they can produce different relative apertures, field-of-view angles, and focal lengths. Reflective telescope systems have low energy-absorption loss, light weight, a large aperture, and no chromatic aberration.

Therefore, reflective telescope systems are typically used as receiving systems. According to their design characteristics, common reflective telescope systems include various types, such as the Newtonian, Cassegrain, Gregorian, and off-axis triple-reflector [5].

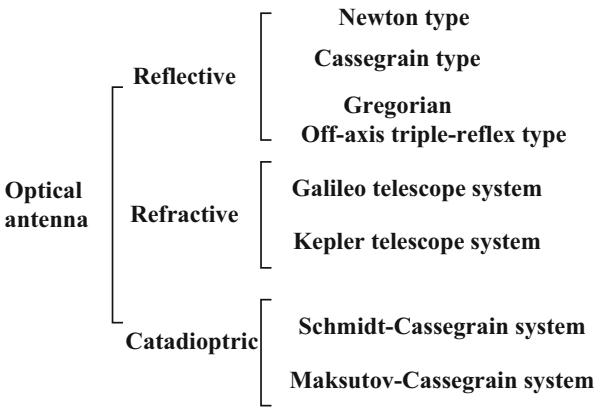


Fig. 11.1 Optical antenna classification [4]

Newtonian Telescopic System

The Newtonian telescope system is shown in Fig. 11.2. It mainly consists of three parts: a parabolic primary mirror, a planar secondary reflector, and a collimating lens [6]. Compared with other telescopic systems, the Newtonian telescopic system has a large relative aperture to adequately receive light and therefore is widely used in optical systems with large apertures. It has a simple structure, no dispersion, and a small spherical aberration.

However, the Newtonian telescope system also has some disadvantages. The structure adopts an open design, and air can circulate freely. This leads to the possibility of airflow interference inside the telescope, thus affecting the image quality. The production cost is high and there is coma for off-axis light.

Cassegrain Telescope System

The Cassegrain-type telescope system is shown in Fig. 11.3. It is a telescopic optical system with a double-reflection structure, consisting of two parts: a parabolic primary mirror and a hyperbolic secondary mirror. In this telescope system, the focuses of the primary and secondary mirrors coincide, thus eliminating spherical aberration and providing high-quality imaging results. In addition, the Cassegrain telescope system is small, compact, and free of chromatic aberration and has an adjustable aperture.

Because the primary and secondary mirrors of the Cassegrain-type telescope system are located on the same optical axis, a certain degree of occlusion exists, caused by the size of the secondary mirror. Consequently, the system field of view is relatively small and the mirror barrel is long [7]. The advantages of Cassegrain-type optical systems, such as compactness and the absence of chromatic aberrations, make them widely used in laser communication, astronomical observation, and remote sensing. Their advantages, such as a short length and light weight, make

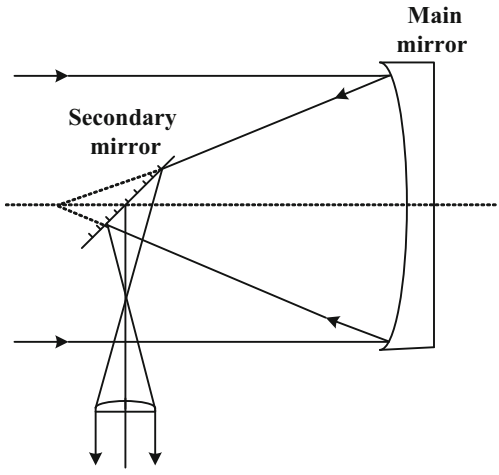


Fig. 11.2 Newtonian telescopic system [6]

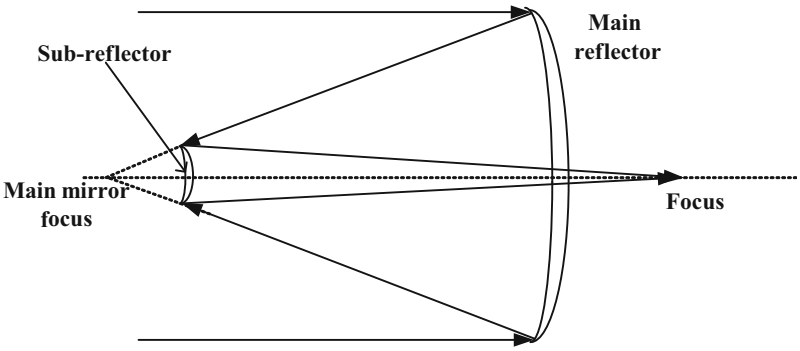
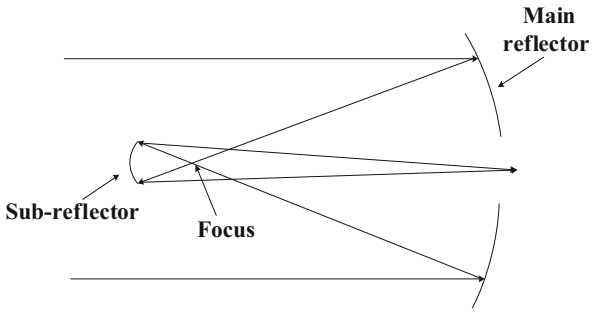


Fig. 11.3 Cassegrain telescope system [11]

Fig. 11.4 Gregorian telescope system [13]



them widely used in the fields of satellites and unmanned aerial vehicles (UAVs) [8–10].

Gregorian Telescope System

The Gregorian-type telescope system is shown in Fig. 11.4. It is a common reflecting telescope, which mainly consists of two parts: a parabolic primary mirror and a concave ellipsoidal secondary mirror. In this telescope system, the focuses of the parabolic primary mirror and concave ellipsoidal secondary mirror coincide, which can eliminate spherical aberrations.

However, the large axial size of the Gregorian-type telescope system can be limiting in some applications and the fabrication process is demanding; hence, there are few practical applications [12]. It is commonly used in astronomical observation, satellite communication, and other fields.

Off-Axis Triple-Reflex Telescope System

The off-axis triple-reflex telescope system is shown in Fig. 11.5. It is a non-coaxial optical system using three aspherical mirrors as the initial structure for off-axis optimization [14]. Its most important features are that it overcomes the chromatic

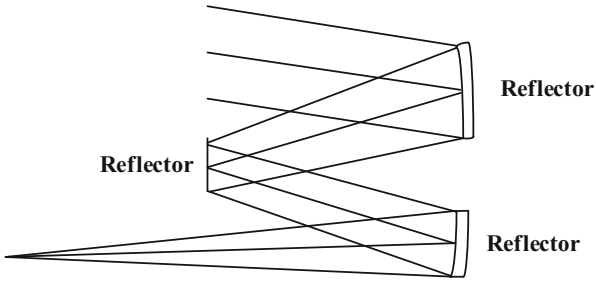


Fig. 11.5 Off-axis triple-reflex telescope system [16]

Table 11.1 Comparison of the advantages and disadvantages of reflective systems

Reflective type	Advantages	Disadvantages
Newton	Simple structure, no dispersion	Open structure, air circulation, maintenance difficulties, high cost, existence of coma
Cassegrain	No chromatic aberration, good image quality, short system length, small size, light weight, compact structure	Small field of view with central occlusion
Gregorian	No spherical aberration	Large axial dimensions, secondary-mirror blocking
Off-axis triple-reflex	Large caliber, long focal length, no chromatic aberration, no central obscuration	Difficult to assemble

aberration and central occlusion of conventional telescopes and has a very high imaging quality.

In addition, the aperture of the off-axis triple-reflex telescope system can be made very large, while its long focal length effectively improves the field of view of the optical system. Owing to the non-coaxial structure, the off-axis triple-reflex telescope system has a very flexible mirror layout, which can be used in many scenarios. It is widely used in the fields of astronomical telescopes, military reconnaissance, and satellite detection. However, its assembly and commissioning are difficult and require more delicate processes and techniques to ensure the optical performance [15] (Table 11.1).

Refractive Telescope Systems

Refractive telescope systems usually consist of an objective lens and an eyepiece. The objective lens converges light onto the focus point and magnifies the image through refraction. The eyepiece receives the image point and magnifies it further to obtain a clearer image [17].

Refractive telescope systems have the advantages of a wide field of view, high contrast, and high clarity. The chromatic aberration present in the system can be addressed by using multiple sets of lenses, but accordingly, it leads to increased cost and loss of luminous flux. In addition, refractive telescope systems are larger, owing to the long length of the lens barrel.

Refractive telescope systems are widely used in astronomical observation, Earth observation, remote sensing, and atmospheric laser communications [17]. They can be used as both transmitting and receiving optical systems. They mainly include Galileo and Kepler telescope systems [18].

Galileo Telescope System

The Galileo telescope system is shown in Fig. 11.6. It is an optical system consisting of a convex-lens objective lens and a concave-lens eyepiece. The objective lens of this system collects the light from the distant object and focuses it at the focal point behind the objective lens, while the eyepiece magnifies the image behind the objective lens and forms a visible virtual image [19].

This system has the advantages of a short distance between the objective lens and eyepiece, a short barrel, aberration elimination, and ortho-imaging. The disadvantages are that the system has a small field of view and no real exit pupil.

Kepler Telescope System

The Kepler telescope system is shown in Fig. 11.7. It mainly consists of an objective lens and an eyepiece, both of which are convex lenses, with a focus point between the objective lens and the eyepiece. To obtain an extended beam that is less affected by the Gaussian-type intensity distribution, it is necessary to include a small-aperture diaphragm in the optical path so that the peak portion of the Gaussian-type intensity distribution passes through [17].

However, this optical antenna still has some drawbacks. For example, a certain amount of light energy is blocked when the small-aperture diaphragm is added, producing a strong light effect. This problem can be solved by adding a spatial filter to distribute the laser energy more uniformly [21].

Table 11.2 compares the advantages and disadvantages of refractive telescope systems.

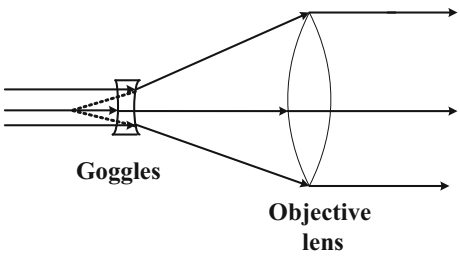


Fig. 11.6 Galileo telescope system [20]

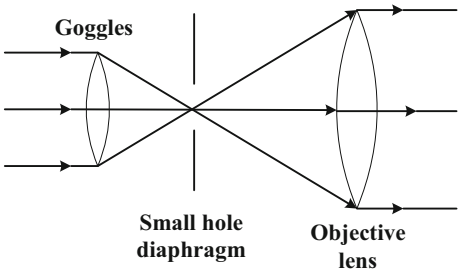


Fig. 11.7 Kepler telescope system [20]

Table 11.2 Comparison of the advantages and disadvantages of refractive systems

Reflective type	Advantages	Disadvantages
Galileo	Short distance between objective lens and eyepiece, short lens barrel, positive image	Smaller field of view and no solid exit pupil
Kepler	A diaphragm can be placed at the focal-point overlap to eliminate stray light effects	Light energy is frequently blocked, strong light effect, relatively long tube length

Catadioptric Telescope System

A Catadioptric telescope system combines refractive and reflective optical systems. To correct the spherical aberration of the primary mirror, the system often uses a wavy correction lens with a convex center, concave side, and complex shape to form the objective lens of the system [17].

The Catadioptric telescope system has many advantages, such as a small optical-system size and high optical performance. It is widely used in the astronomy, military, and aviation fields. The Schmidt–Cassegrain and Maksutov–Cassegrain systems are two common examples of reflexive optical systems [22].

Schmidt–Cassegrain System

The Schmidt–Cassegrain system is shown in Fig. 11.8. It consists of two reflecting surfaces, a concave parabolic primary mirror, and a convex hyperbolic secondary mirror [23]. The Schmidt–Cassegrain system has many deformed structures, such as a double spherical-mirror set structure, double aspheric-lens set structure, and combined spherical and aspheric mirror structure [17]. The primary reflector has a shorter focal length, which is more compact than a conventional parabolic reflector, and can obtain a much higher gain than other types. However, the system suffers from problems, such as secondary-mirror occlusion and the difficulty of making large-aperture Schmidt correction plates.

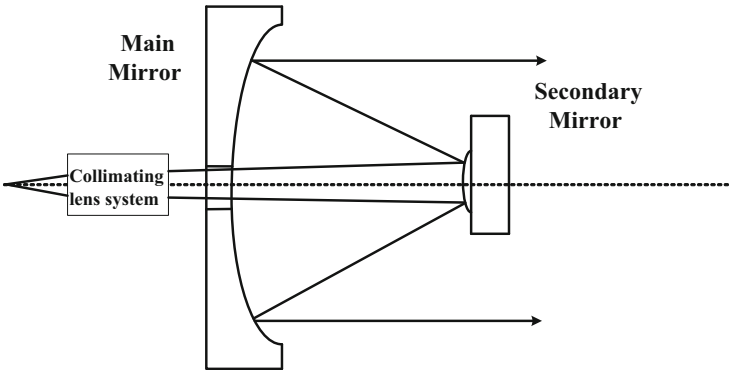


Fig. 11.8 Schmidt–Cassegrain system [23]

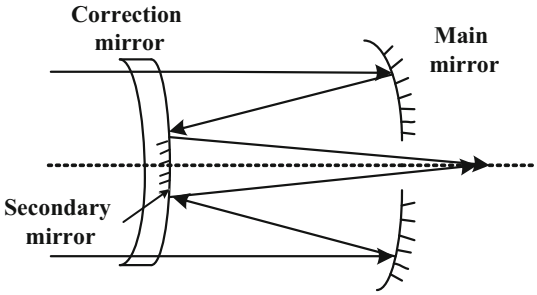


Fig. 11.9 Maksutov–Cassegrain system [25]

Maksutov–Cassegrain System

The Maksutov–Cassegrain (Mak-Cass) system is shown in Fig. 11.9. It employs a large-curvature meniscus-type correction lens and an aluminized dot on the correction lens as a secondary reflector [24]. Moreover, it adds an aspherical lens to the Schmidt–Cassegrain system. After the light passes through the primary and compensating lenses, it passes through the aspherical lens for compensation to achieve higher optical performance. When the Mak-Cass antenna is used as a spatial optical-coupling system, it has large-aperture coupled reception and an improved ability to converge optical power [17].

A typical feature of the Mak-Cass telescope system is the use of a fold-and-reverse structure, which makes its subreflector very small and more compact. This greatly reduces the size and weight of the antenna and enables large-aperture and long-focal-length imaging [25]. However, the telescope system suffers from obscuration problems, leading to a large loss of energy. Mak-Cass telescope systems are widely used in the military, industrial, and aeronautical fields.

Table 11.3 compares the advantages and disadvantages of Catadioptric telescope system.

Table 11.3 Comparison of the advantages and disadvantages of reflexive systems

Reflexive type	Advantages	Disadvantages
Schmidt–Cassegrain	Compact system for higher gain, excellent image quality, no coma or dispersion	Poor ball presence, large volume, high cost
Maksutov–Cassegrain	Compact, long focal length, large aperture	Large volume, small field of view

Problems of Point-to-Point Antennas

Traditional optical wireless communication research and experiments typically use point-to-point antennas, which face the following problems.

1. Distance: Generally, point-to-point antennas must be applied in closer locations. In more distant locations, the communication quality is reduced.
2. Weather influence: Optical transmissions are susceptible to atmospheric influence, such as fog, rain, haze, etc. Such weather conditions can cause a large amount of optical-signal loss, which, in turn, affects the communication quality. Bad weather can also make deployment and maintenance more difficult.
3. Line-of-sight obstacles: Because light propagates in a straight line, it cannot bypass obstacles in the communication path. It must have a clear line of sight between the two sides of the communication, limiting its use and scenarios.
4. High interference: Light signals are susceptible to daylight and other light sources and are also susceptible to electromagnetic waves.
5. High cost: Because of the special devices and optical components required for optical wireless communication, it is more expensive.

Foreign and Domestic Research Progress

In 1953, the United States officially launched the Star Wars program, which promoted the rapid development of space laser communication. Since then, Germany, the United Kingdom, Japan, and other countries have also been involved in this field and started related work [26].

In 1993, Clarke et al. [27] proposed the use of a large field-of-view optical antenna to achieve all-round coverage of near-Earth targets with a small movable detector on its focal plane to realize point-to-multipoint optical communication.

In 2002, the Transformational Satellite Communications System (TSAT) [28] proposed to place multiple receiving terminals at different locations to achieve simultaneous transmissions from multiple locations.

In 2002, Sherman M P et al. [29] proposed an optical wireless communication system, as shown in Fig. 11.10. The telescope in this system had an aspherical mirror design, such as the Ritchey–Chrétien (RC) telescope shown in (a). It featured a concave primary mirror and convex secondary mirror, both of which had a hyperbolic shape. The mirror had a large focal plane, which allowed the automatic

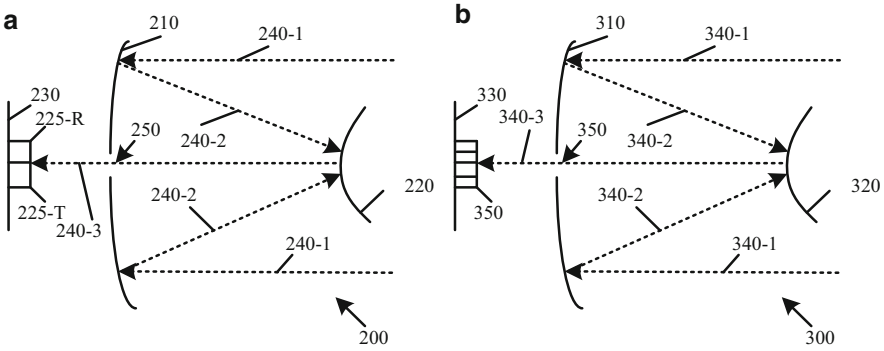


Fig. 11.10 Optical Wireless communication system [29]·(**a**) RC optical telescope (**b**) Point-to-multipoint communication system

matching of fixed or stationary mirrors between transmitter and receiver, thus greatly reducing production costs. An $N \times N$ fiber array could be established on the focal plane of the RC optical telescope for point-to-multipoint communication, as shown in (b).

In 2006, Zhou L et al. [30] proposed a new optical-antenna array for interstellar laser communication and performed a statistical analysis of the causes of the degradation of its alignment-error performance to verify the feasibility of the new optical-antenna array. Compared with ordinary antennas, it had the advantages of a small size and adaptive pointing.

In 2008, Zhang C et al. [31] proposed a novel multipoint ground-reception scheme to improve the communication quality. The problems of communication-link interruption and increased BER were mitigated by channel redundancy.

In 2008, Zhang W L et al. [32] proposed a theoretical structure for one-to-many and many-to-one synchronous communication using semiconductor lasers.

In 2011, Li K et al. [33] proposed a one-to-many network communication scheme using the sea surface as a laser diffuse-scattering medium. They applied the Kirchhoff approximation theory to analyze the transmission characteristics of ocean diffuse communication.

In 2011, Jiang H L et al. [34] proposed an optical antenna suitable for simultaneous multipoint duplex communication. The working principle and implementation method of an optical antenna with three concentric spheres in laser-communication networking were studied. Through theoretical analysis and experiments, the acquisition, pointing, and tracking (APT) control model of the multi-reflector splicing from optical transceiver was studied to help lay the foundation for the application of space laser-communication-link networking.

In 2012, Jiang H L et al. [35] proposed a multi-target inter-synchronous laser-communication technology and applied it to the key technologies of optical antennas, relay optical systems, transceivers, and APT. It was based on a rotating paraboloid with a multiple-reflector splicing structure to provide a new technical idea for optical wireless communication-link networking.

In 2012, He Y J et al. [36] designed a multipoint bi-directional laser-communication optical antenna based on a splicing mirror. The antenna broke through the structural limitations of traditional antennas and could be applied to space laser-communication networks as node terminals.

In 2013, Sun X D [37] proposed an eccentric cross-gimbal coarse-tracking structure scheme, which used a rotating paraboloid as the base and a spliced mirror. By studying the construction of a one-to-many communication optical terminal, a servo control unit was designed for the system to achieve accurate alignment and tracking functions.

In 2013, Xiao P [38] proposed a new all-optical capture technique for one-to-many laser-communication systems and discussed the capture method of this technique in depth. The parameters and capture methods of the all-optical capture technique were analyzed in detail, and a model was established according to the capture strategy.

In 2013, Yu X N et al. [39] established a servo control system for capture, targeting, and tracking to realize a one-to-many space laser-communication network.

In 2014, Yu X N et al. [40] proposed a single-beam tracking system for laser-communication networking, which could accurately track the spot position of interstellar laser communication. It laid the foundation for the next stage of realizing many-to-many laser-communication networking.

In 2014, Jiang H L et al. [41] proposed a simultaneous dynamic-communication technology between multiple points, which provided technical support for the formation of an integrated information network between heaven and earth.

In 2014, Meng L X et al. [42] proposed a general idea about optical antennas to simultaneously track multiple targets and established an information-transmission network. The kinematic model was simulated to obtain the ideal field of view of the optical antenna, which provided a theoretical basis for subsequent development.

In 2014, Meng L X et al. [43] proposed a scheme using a multi-mirror reflector, with a rotating paraboloid as the substrate, as an optical antenna to establish a one-to-many laser-communication network. The system adopted an eccentric cross-tracking frame design, and the problems of decreasing optical-energy utilization and system stability were effectively solved.

In 2014, Meng L X [44] proposed a rotating paraboloid-based multi-mirror stitching capture and tracking system with dynamic-optics principles to establish tracking and scanning models. Experimental results showed that this system could rapidly capture and accurately track multiple targets with an azimuth angle of 360° and a pitch angle of ±15°.

In 2014, Zheng Y et al. [45] proposed a four-reflector optical antenna structure and a link-satellite networking scheme based on the optical antenna structure and conducted an optical dynamic study on the antenna. This antenna structure was capable of laser-source capture within a 50° field of view and a 40° pitch angle. This scheme achieved link-satellite networking and improved the communication efficiency.

In 2014, Zhao K H et al. [46] proposed an optical-antenna structure capable of simultaneously tracking and communicating with multiple targets. This is the core

structure of one-to-many laser-communication networking technology, and the tracking performance is the key factor for measuring its technical reliability. To verify the tracking performance of the optical antenna, the priority was the design and development of an optical-antenna tracking-performance test system.

In 2014, Zhang L J et al. [47] proposed the use of Kalman-filter prediction to track an optical spot. Simulation results showed that the Kalman filter could well predict the velocity and motion direction of a moving object within the specified error. This could be used to control the spot position of the optical lens and provide a theoretical basis for realizing multi-point communication.

In 2014, Zhang L J [48] proposed a multi-point synchronous tracking technique for satellite space laser-communication networking. A research method combining theory and experimental tests was used to design an overall scheme for one-to-multipoint laser communication.

In 2015, Zhao K H et al. [49] designed a complete test system to verify the tracking performance of an optical antenna to ensure the reliability of one-to-many laser-communication networking technology. The tracking performance was tested by providing a laser signal to the optical antenna as the target source. In the actual communication, the optical antenna tracked the moving target reliably and ensured the communication quality of one-to-many laser communication.

In 2015, Yu X N et al. [50] discussed the design and development process of servo antennas in one-to-many laser-communication networks and conducted splicing and tracking experiments. The success rate of the splicing reliability reached 90% and the tracking error was less than 50 μrad, which guaranteed the reliable operation of the communication networking system.

In 2015, Jiang H L et al. [51] proposed an overall optical scheme for a pair of simultaneous point laser communications, deduced the workflow of the APT system, and analyzed the relationship between the off-target amount and the rotation of the reflector, during the capture alignment and coarse tracking.

In 2015, Jiang L et al. [52] proposed a “one-to-two” laser-communication networking technique and also calculated the power margin of the interplanetary laser-communication link. The experimental results showed that the technique could achieve a 200-km communication distance and 2.5-Gbit/s communication rate, providing a new technical approach for intra-flight communication networking in formations.

In 2015, Zhang T et al. [53] proposed a new optical-antenna design method, which was capable of one-to-many laser communication. They derived the operating range of multiple optical antennas and their inter-platform communication, providing a feasible design solution to achieve multi-point inter-communication.

In 2015, Ke X Z et al. [54] designed a transceiver optical antenna based on the Cassegrain telescope. It used the effective area as the transmitting and receiving area of the antenna to separate the optical path, thus improving the transmitting efficiency of the antenna.

In 2016, Zhang T [55] proposed several design schemes for laser-communication grouping in antenna systems. Among them, a new theory of optical properties was derived, based on the rotating-paraboloid antenna, which achieved greater dynamic

communication within an azimuth of 360°, as well as two-way communication between multiple platforms.

In 2016, Zhang T et al. [56] proposed a novel laser-communication networking system that enabled efficient laser communication by precisely focusing the laser through a confocal reflection technique. The system could be used for networked communication between different platforms and had various advantages, such as a simple structure, light weight, and easy miniaturization.

In 2016, Ma G D et al. [57] investigated and modeled the tracking-antenna performance in one-to-many laser communication using a joint machine–electric simulation. The simulation results showed that the performance and error of the tracking antenna could be obtained accurately and quickly after the joint-simulation approach, and the servo-drive error was also reduced.

In 2016, Jiang L et al. [58] proposed to use the combination of a wide-angle beam expander and dual optical-wedge set as an optical antenna for a multipoint laser-communication system. This antenna expanded the beam angle and eliminated the optical offset; it could be used for dynamic tracking and signal reception of the target. This optical antenna design provided a new technical method for space laser-communication networking technology.

In 2017, Zhang T et al. [59] proposed two different schemes to solve the problem of communication-platform networking: one based on the rotating paraboloid model and another for a laser-communication networking antenna system.

In 2017, Zhao Y et al. [60] proposed an underwater optical wireless link based on orbital angular momentum (OAM) to solve the problem of underwater one-to-many data-information transmissions. This technology had a high speed, low power consumption, and interference resistance.

In 2019, Li X et al. [61] proposed an optical-antenna structure with high carrying capacity, and the precise design of the connection of various optical elements ensured the accuracy of the optical antenna position. The experimental results showed that the imaging quality after the bearing was significantly improved, compared with that before the bearing, which helped lay the foundation for one-to-many laser communication.

In 2019, LOU N [62] proposed a wireless laser-communication abutment design, which could carry out multipoint laser communication to multiple bands. To effectively transmit and receive laser communication, a complete set of laser-communication equipment containing a relay system and splitting system was designed.

In 2020, Cao H Y [63] proposed a new networking scheme, which used three independent single-reflector-type follow-up units to adjust the optical axis to ensure the optical system for accurate targeting. It provided a strong support for the development of one-to-many laser-communication networking technology.

In 2020, Wang S D [64] designed a bearable optical antenna that could efficiently transmit data and could be used for inter-satellite platform communication. The optical antenna was designed using an orthogonal test optimization method for the main mirror to achieve high-quality imaging and a stable carrying capacity.

In 2020, Song G et al. [65] proposed that an optical frequency comb (OFC) could be used instead of a laser or laser array as a carrier source for a pair of multiple optical wireless communication systems.

In 2020, Wang J Y et al. [66] proposed a dual-mirror linkage-control scheme for reflector-spliced optical antennas for multi-target tracking. The optical-signal energy utilization was improved. Meanwhile, the reflector-spliced optical antenna of this technique achieved efficient and accurate target capture and real-time multi-target tracking, which greatly improved the performance and reliability of the optical detection system.

In 2021, Wang J Y et al. [67] designed a multi-target APT system, based on an optical antenna and fast-steering mirror (FSM), and used a mirror-linked control method to improve the utilization efficiency of the optical antenna. A principle prototype was developed and dynamic demonstration experiments were conducted for point-to-multipoint laser communication.

In 2021, Wang T et al. [68] designed a "one-to-two" space laser-communication system to realize high-speed laser data transmissions of multiple nodes in space information networks.

In 2021, Bai Y Y et al. [69] explored a "one-to-four" space laser-communication system, elaborated its working principle, and studied in depth the networking and control mode of the system. In this system, the optical-transceiver design relied on a mathematical model based on the Jetlink control technology. It provided an effective solution to communicate between multiple points in space laser communication.

In 2021, Fan T et al. [70] described the principle of a networked "one-to-four" main optical-transceiver system. By designing a fuzzy self-anti-disturbance control method, the precise tracking control of the main optical transceiver was realized, which greatly improved the stability and reliability of the laser-communication system.

In 2021, Li X M et al. [71] conducted a theoretical analysis of the support parameters of an integrated SiC/Al pendulum mirror to reduce the impact of planarity errors on the reflector surface shape. By optimizing the design, the pendulum-mirror surface-shape accuracy requirements for one-to-many laser-communication networks were met. In the same year, Li et al. [72] designed an integrated compact pendulum mirror that could realize one-to-many space laser communication in a large range. Higher requirements were imposed on the structural dimensions of its components to improve the reflection efficiency of the pendulum mirror.

In 2021, Zhao B Q et al. [73] studied the servo performance of the reflector optical antenna in a one-to-many laser-communication network and simulated and analyzed the optical closed-loop performance of the velocity and position loops. The simulation results showed that the optical closed-loop strategy of the velocity loop had a larger error in the tracking computation compared with the optical closed-loop strategy of the position loop. However, its servo bandwidth was larger than that of the position loop and had a stronger anti-disturbance capability.

In 2021, Xu M et al. [74] proposed a pair of multiple laser-communication optics, based on which, a principle prototype was designed and fabricated and experiments were conducted. The performance results showed that the principle prototype had

good multidimensional detection performance. A principle prototype of integrated multi-dimensional space-target detection with laser communication with a 250-mm optical aperture was developed.

In 2022, Zhang K Q et al. [75] established a corresponding coarse-tracking system model for a pair of multi-network optical-communication coarse-tracking servo control systems and designed a model-predictive controller (MPC) for this model to effectively improve the system tracking-control performance.

In 2023, Ding C et al. [76] used a pulse-width modulation technique combined with the data-routing and forwarding function of a laser-relay station for optical wireless communication networking. They demodulated the received signal using a high-speed digital-to-analog converter to improve the stability of the optical wireless communication system. The system was more flexible and had higher communication quality than the point-to-point laser-communication system, while the system achieved one-to-many optical wireless communication without adjusting the transmission angle.

In 2023, Hu Y B et al. [77] established a two-way tracking and targeting model for satellite laser-communication optical terminals to realize "one-to-four" network communication. The concept of equivalent composite control was introduced to further improve the tracking accuracy without changing the parameters and stability of the original calibration-control system.

In 2023, Wang L H et al. [78] designed a scheme to simultaneously stabilize multiple optical axes using a single gyroscope. They proposed a method to solve the attitude of the axes for the miniaturization, weight reduction, and networking of a simultaneous point-to-multipoint laser-communication optical transceiver.

One-point-to-multipoint optical wireless communications have their own advantages and disadvantages between the various structures, according to the integration of their respective characteristics that are categorized as shown in Table 11.4.

One-to-Many Optical Communication

At present, research on optical wireless communication systems is mainly focused on single-point-to-single-point communication structures. Although this structure form can meet the demand in some scenarios, in practical applications, multiple terminals will need to communicate with each other. In this case, the traditional single-point-to-single-point transmission form can no longer meet the communication needs. It must establish a complete communication network for efficient information transmission through a pair of multi-laser communication systems to achieve a higher practical value.

Basic Principles of One-to-Many Optical Communication

In multi-target communications, the visual range of the optical antenna is critical. Only with a sufficiently wide visual range can a system communicate with multiple

Table 11.4 Research progress in one-to-multipoint optical wireless communication

Year	Researchers	Key technologies	Study results
1993	Clarke	Large field-of-view optical antenna	Smaller field of view, serious stray light, difficult device arrangement
2002	TSAT	Multi-transceiver combined optical antenna	
2002	Sherman	Optical wireless communication system	
2008	Zhang	New multi-point terrestrial-reception solution	Mitigates communication-link disruptions and increases BER
2011	Jiang H L	Optical antenna for simultaneous multi-point duplex communication	Simple structure and wide communication range, long communication distance
2014	Zheng Y	Four-reflector optical-antenna structure	
2016	Jiang L	Wide-angle beam expander and dual optical-wedge set combined with an optical antenna	
2013	Sun X D	Eccentric cross universal frame coarse-tracking structure scheme	Enables precise alignment and tracking functions
2014	Yu X N	Single-beam tracking system	
2014	Meng L X	Multi-reflector stitching capture and tracking system	
2020	Wang J Y	Multi-target tracking for optical antennas	
2015	Zhang	New optical-antenna design method	Provides a viable design solution for inter-multipoint communication
2015	Ke X Z	Integrated optical antenna for transmitting and receiving	Separation of transceiver and transmitter to improve transmitting efficiency
2011	Li K	One-to-many network communication solution	New technical idea for one-to-many space laser-communication networking
2012	Jiang H L	Multi-reflector structure as a shared transceiver antenna	
2012	He Y J	Multi-point bi-directional laser-communication optical antenna	
2020	Wang S D	Bearable optical antenna	
2021	Xu M	One-to-many laser-communication optical principle	
2023	Wang L H	View-axis attitude solution method	
2014	Meng L X	One-to-many laser-communication network	High communication quality, high stability, high imaging quality
2019	Li X	Optical antenna with high load-carrying capacity	
2021	Wang T	“One-to-two” space laser-communication system	
2021	Bai	“One-to-four” space laser-communication system	
2023	Ding	Optical wireless communication networking	

targets simultaneously without compromising the communication quality. Current-optical-communication systems commonly use independent-aperture optical systems, which cannot communicate with multiple targets simultaneously [79].

The challenge of simultaneous communication to multiple targets can be solved if an optical system can be designed such that incident light from any angle in space (over a wide range) can pass through the system and be focused on the same point. The design of such an optical system requires full consideration of optical principles and relevant mathematical models, as well as detailed design and testing, to ensure its efficiency and reliability in practical applications.

By comparing the optical properties of various surfaces, it was found that the rotating paraboloid had the unique property that all reflected beams were parallel to the rotation axis, after the incident light was reflected by the rotating paraboloid. It could be used to design various optical devices, such as optical antennas, reflectors, etc. The optical principle is shown in Fig. 11.11.

As shown in Fig. 11.11, assume that the equation of the parabola is

$$y^2 = 4px. \tag{11.1}$$

$P(x_0, y_0)$ is a point on the parabola. If we apply it to (11.1), $y_0^2 = 4px_0$. The axis of symmetry of this parabola is the X-axis; that is, $y = 0$. The tangent line through point p is

$$y_0 y = 2p(x + x_0) \rightarrow 2px - y_0 y + 2px_0 = 0. \tag{11.2}$$

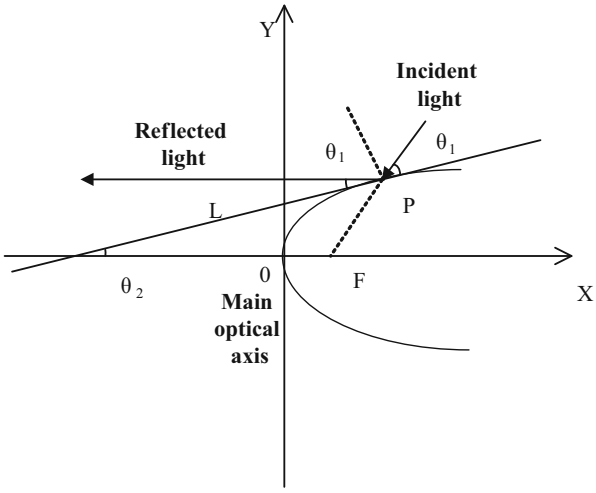


Fig. 11.11 Schematic diagram of the parabolic optics principle [44]

Point F is the focus of the rotating paraboloid with coordinates $(p, 0)$, and the equation of line PF is

$$y = \frac{y_0}{x_0 - p}(x - p) \rightarrow y_0 x - (x_0 - p)y - py_0 = 0. \tag{11.3}$$

With the basic principle of reflection, equations can be obtained:

$$\begin{aligned} \cos\theta_1 &= \frac{|2py_0 + (x_0 - p)y_0|}{\sqrt{4p^2 + {y_0}^2}\sqrt{{y_0}^2 + (x_0 - p)}} \\ &= \frac{|x_0 y_0 + py_0|}{\sqrt{4p^2 + 4px_0}\sqrt{4px_0 + {x_0}^2 - 2px_0 + p^2)}} \\ &= \frac{|y_0(x_0 + p)|}{\sqrt{4p(p + x_0)}\sqrt{(x_0 + p)^2}} \\ &= \frac{|y_0|}{\sqrt{4p(p + x_0)}} \end{aligned} \tag{11.4}$$

$$\cos\theta_2 = \frac{|-y_0|}{\sqrt{4p^2 + {y_0}^2}} = \frac{|-y_0|}{\sqrt{4p^2 + 4p{x_0}^2}} = \frac{|y_0|}{\sqrt{4p(p + x_0)}}. \tag{11.5}$$

From Eqs. (11.4) and (11.5), we obtain $\cos\theta_1 = \cos\theta_2$. Because θ_1 and θ_2 are both acute angles, $\theta_1 = \theta_2$. Assuming that the incident light is a straight line, PF, there exists a reflected light parallel to the axis of symmetry. By exploiting the optical properties of the paraboloid, the receiving antenna can receive light beams from different directions, thus enabling the simultaneous tracking and communication of multiple targets. Figure 11.12 illustrates the working principle of a pair of multiple optical-communication optical systems.

The system is divided into three parts: the optical antenna, relay optics, and splitting system [35]. In the optical wireless communication system, the optical-antenna part undertakes the task of signal acquisition and transmission. When multiple optical signals arrive at the optical antenna, they pass through a multi-reflector splicing and servo system so that the light is incident along the optical axis into the relay-optics system. The optical-antenna part generally adopts structures, such as a Cassegrain telescope system, off-axis two-inverse telescope system, and refractive system, to achieve a common laser-transceiver aperture [80].

The relay-optics part is mainly responsible for converging the received communication light and expanding the transmitted beacon and communication light. The splitting-system part is mainly responsible for separating the received beams for separate processing and directing them to the corresponding detectors for detection, thus realizing multi-point laser communication [44].

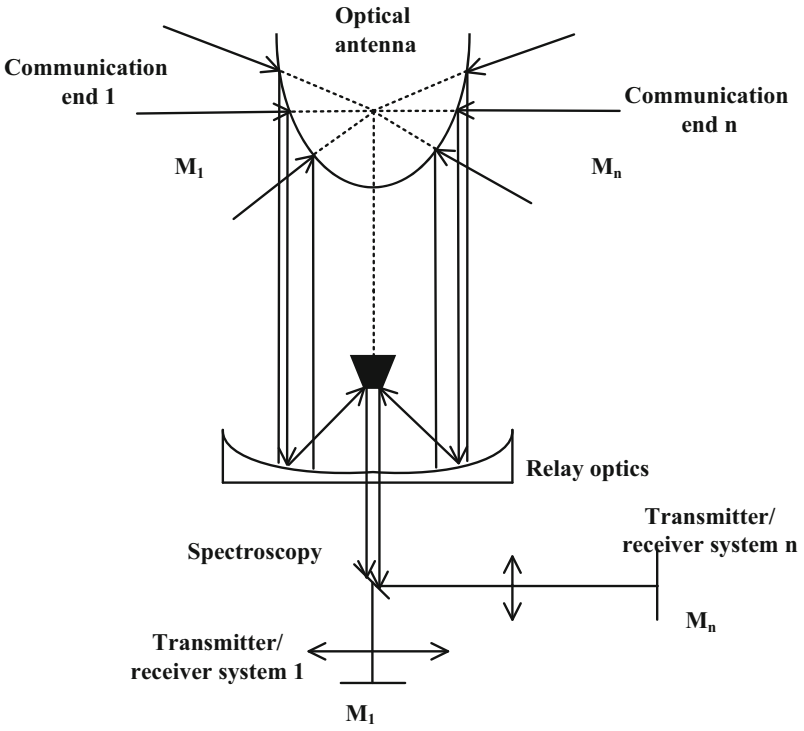


Fig. 11.12 One-to-many optical-communication system working-principle sketch [44]

Acquisition, Pointing, and Tracking (APT)

In space laser communication, owing to the small beam-dispersion angle and long transmission distance, precise tracking techniques must be adopted to ensure the transmission and reception quality of the laser signal. Here, APT technology becomes particularly critical and directly affects the quality of the space laser communication [81].

1. Acquisition: When the beacon light from the transmitter is received, the beam is usually focused on the center of the detector to identify the target. To identify the target in an uncertain area, a scanning operation is required to find and capture the target's beacon-light signal, thus laying the foundation for subsequent targeting and tracking [82].
2. Pointing: After capturing the beacon-light signal, it must be transmitted through the receiver. To meet the communication requirements, the transmitter and receiver should be aligned with each other so that the optical axis is consistent.
3. Tracking: Owing to the relative motion between the communicating and receiving ends, a precise connection must always be maintained to ensure that communication is maintained between satellites or between satellites and ground stations. Because of the relative motion between the communicating and

receiving ends, together with the influence of other factors, such as spot flicker and platform vibration caused by atmospheric turbulence, the system must control the error by tracking the aiming after it is completed to ensure that the accuracy is within the allowed range [82].

In optical communication, the signal-transmission loss results in a narrow communication beam. To ensure that the signal energy is accurately received by the receiver, the communication beam needs to be precisely controlled so that it is aligned with the system optical axis and remains within the error range. In this process, APT technology plays an important role. It plays a crucial role in interplanetary laser communication because it has very high precision and therefore becomes an essential technical means to ensure the success of communication [83].

In a one-to-many space laser-communication system, the workflow of the APT system consists of five main stages, namely,

1. Guidance pointing
2. Capture alignment
3. Coarse tracking
4. Fine tracking
5. Communication [84]

These are performed sequentially to achieve high-precision tracking and communication for one-to-many laser communication.

The main task of guidance pointing is to clarify the objectives and requirements of inter-multipoint communication and to determine information about the target. The main task of capture alignment is to determine the corresponding working reflector, based on target information, and achieve capture alignment so that the beam can be incident parallel to the subsequent optical system.

The main task of coarse tracking is to track a large field of view, based on the capture alignment, and to perform closed-loop tracking so that the beam is successfully incident to the center of the field of view [85]. The main task of fine tracking is to track the target on the basis of coarse tracking and suppress the residual error from coarse tracking so that it meets the communication requirements.

Coarse tracking requires a wide search range and fast target locking, while precision tracking requires high accuracy, a smaller search range, and advance calculation. The main task of communication is to achieve two-way communication with stable tracking [52].

Optical Principle of One-to-Many Communication

According to the one-to-many optical principle, terminals can be mainly classified into the following types: multi-optical terminal combination, multi-reflector, TSAT grouping, dual optical-wedge grouping-combination, fiber-array structure, multi-transmitting and multi-receiving structure, three-concentric-sphere structure, and rotating-paraboloid structure [86].

Multi-optical Terminal Combination Type

The structure of the multi-optical terminal combination type is shown in Fig. 11.13. It uses a single optical transceiver to realize one-to-many laser communication by combining multiple optical transceivers [87]. The optical transceiver mainly consists of an APT system, communication transceiver system, and optical subsystem.

The APT system consists of coarse and fine composite axes, in which the coarse-tracking structure is realized by the reflector, while the fine-tracking structure is realized by a piezoelectric ceramic oscillator, and the coarse and fine detection share the same CCD camera to realize variable resolution and variable frame-rate control for fast capture and high-precision tracking.

The communication transceiver system includes communication transceiver and receiver units with a bi-directional-rate asymmetry function. Optical subsystems usually include multiple units, such as shared Cassegrain telescope units, optical relays, and two-color beamsplitters [48]. These units work together to enable efficient and reliable signal transmission and processing in optical-communication systems.

By combining multiple optical-terminal units, the communication coverage and capability can be increased, making multi-target communication more convenient and efficient. However, the weight limitation of optical transceivers means that multiple combined optical transceivers must occupy extensive space and resources, increasing the cost and complexity of the system. In practice, they are mostly used in large-platform laser communication.

Multi-reflector Type

The multi-reflector type is shown in Fig. 11.14 and consists of multiple reflectors and a main optical transceiver. The multi-reflector type is innovative technology for space laser-communication networking. Multiple mirrors are placed at different

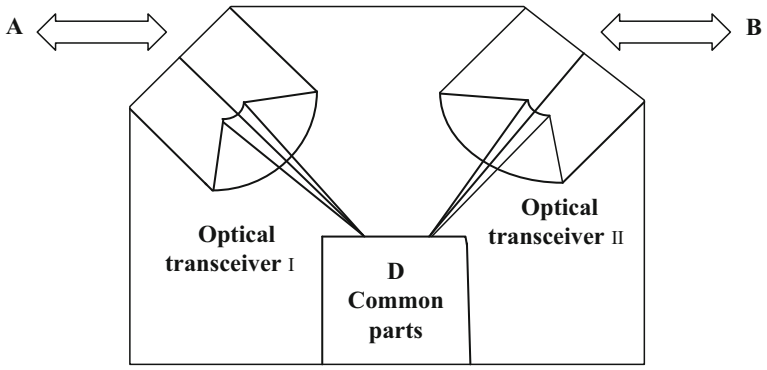


Fig. 11.13 Multi-optical terminal system diagram [87]

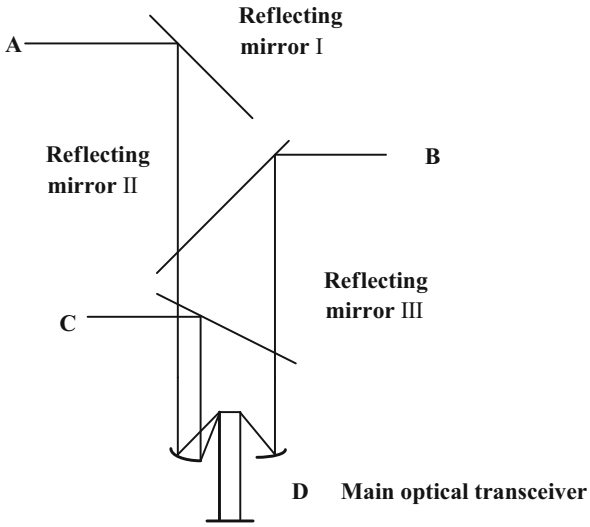


Fig. 11.14 Schematic diagram of a multi-reflector structure system [86]

angles at the front of the optical antenna of the optical transceiver to communicate in different orientations [45].

In the multi-reflector type, each reflector is single-point-to-single-point, and there are no redundant optical components. It has better flexibility and scalability, compared with the traditional optical-communication network built with multiple optical terminals, but it can only be used for "one-to-three" or "one-to-four" laser communications. The number of communications is still limited by the energy utilization and the difficulty of plating the reflector-film system [45]. The multi-reflector type can realize high-speed and high-bandwidth optical communication and has broad application prospects.

TSAT Networking

In 2016, the United States launched the Transformational Satellite Communications System (TSAT), a networking program designed to provide global military high-bandwidth communications and secure high-speed data-transmission capabilities [88].

The TSAT space segment consists of five satellites, of which the laser link is the backbone, integrating the geostationary satellite system, Optical Relay Communications Architecture (ORCA), Advanced Polar System (APS), and Tracking and Data Relay Satellite System (TDRSS). They cover the area between the 65° north and south latitudes and are the next generation of secure broadband satellite communication systems [89].

Compared with traditional microwave communication, the TSAT network introduces inter-satellite laser links. The communication bandwidth is improved and is

expected to reach a communication rate of 10–40 Gbit/s. The scheme covers a wide range of communication means for interplanetary and interstellar applications [90]. Its main advantages are a fast communication speed, high bandwidth, and wide coverage, while various communication means complement each other and play an important role in improving military communication and intelligence exchange, etc. The TSAT networking plan provides new ideas and technical support for the development of the satellite-communication field and has broad application prospects.

Double Optical-Wedge Group Combined Type

The combined double optical-wedge group structure is shown in Fig. 11.15. It uses a combination of a wide-angle beam expander and a double optical-wedge group as the optical antenna to achieve dynamic tracking and signal reception of the target [58]. A wide-angle beam expander is an optical component used to receive the signal light and transmit it to different communication targets. Although it has the advantage of a large field of view, it also has difficulty dynamically tracking moving communication targets.

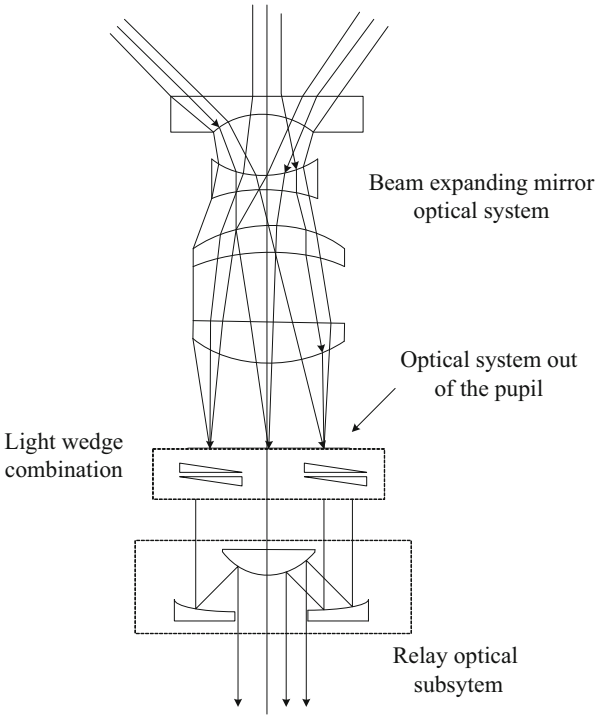


Fig. 11.15 Schematic diagram of the combined optical structure of the double optical-wedge group [58]

To address this problem, a double optical-wedge set is placed at the exit pupil of the optical system. In addition, the system uses a relay-optics subsystem to focus and optimize the signal light [91]. With this structural modification, the system can simultaneously track and receive signals from multiple targets, thus improving the application range of the system.

The combined double-wedge system can control the angle of the outgoing beam and improve the light-energy utilization by adjusting the rotation angle of the double wedge, according to the number and position of the targets. In addition, it has a low cost, has simple structure, and can dynamically track and receive signals from multiple targets. It is widely used in the field of multi-target laser-communication systems, such as communication between aircraft and aircraft and between artillery positions, and has broad application prospects [58].

Fiber-Array Type

The structure of the fiber-optic array is shown in Fig. 11.16. It consists of three parts: the overall optical path, the fiber end face, and the transmitter–receiver end. The overall optical path consists of a transmitting telescope and a receiving telescope. The fiber end face is an $N \times N$ fiber array. The transmitter end is an RC optical telescope that can be used for optical wireless systems, and the receiver end is a point-to-multipoint communication system.

In 2002, Lucent Technologies, Inc. (USA) [92] provided an improved optical telescope for optical wireless communication systems, which had the advantages of high bandwidth, wide link range, and high reliability, compared with the conventional design. The structure used a cassette system that automatically aligned the transmitter and receiver, positioning the $N \times N$ fiber array in the focal plane of the RC optical telescope, thus enabling point-to-multipoint communication with a single optical telescope [93].

Although this structure successfully expanded the field of view, the limitations of the card-based system and the numerical aperture of the fiber resulted in a still smaller field of view, more-severe stray light, and more-difficult transceiver arrangement; it needed to be further optimized [86]. In addition, interactions and interference that occur between fibers when transmitting through fiber arrays need to be overcome. Therefore, future fiber-array technology needs to be further improved and optimized to enhance its transmission efficiency and stability.

Multi-transmission, Multi-receiver Structure

A one-to-many laser communication system proposed by TeraBeam, USA, is shown in Fig. 11.17 [94]. In this system, light source 210 and light receiver 520 are optically coupled by optical input 240 and transmit the light to fiber 270. Optical input 240 is coupled to transmit the light received from light source 210 to reflective surface 230b, which reflects emitted light 260 in the form of a beam, 261, to the corresponding remote location.

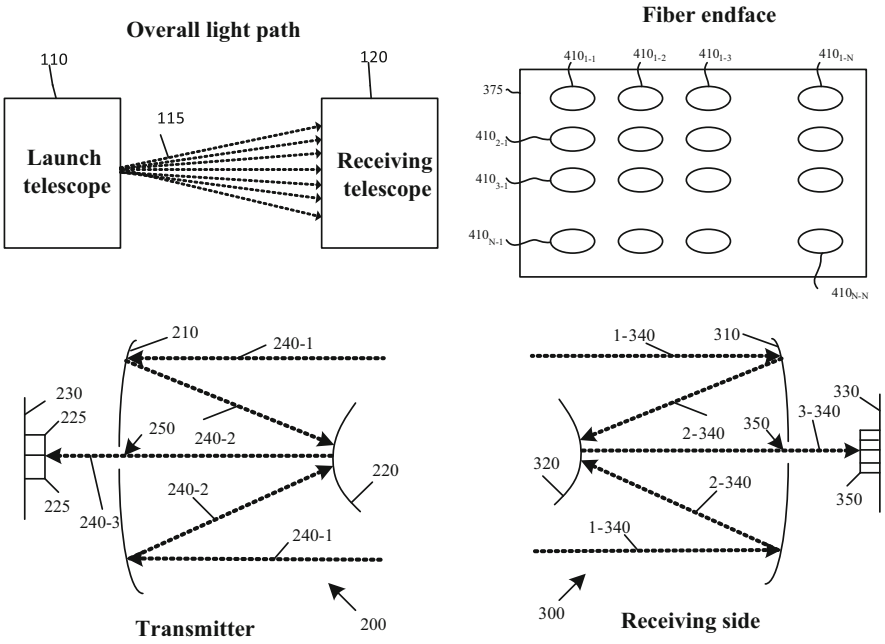


Fig. 11.16 Fiber-array type structure [92]

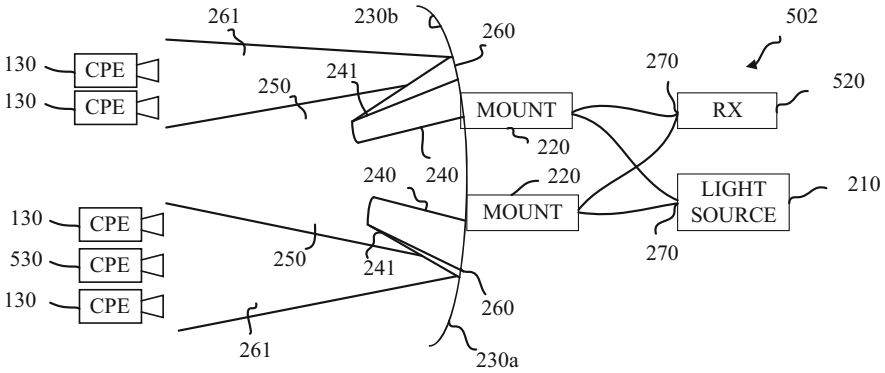


Fig. 11.17 Schematic diagram of one-to-many laser communication in free space [94]

Customer Premise Equipment (CPE) 130 at the remote position receives light beam 261 and returns the light signal back to reflecting element 230a, which is reflected to terminal 240 of 241. When the light signal is returned, it is coupled through optical receiver 520 to optical input 240; it is demodulated by optical receiver 520, converted into an electrical signal, and made available to the connected device.

A technique for tracking movement over long distances is provided by adjusting the position of node unit 250 and can be received simultaneously using optical receiver 520, thus enabling the design of a one-to-many laser-communication networking structure.

Although the structure can expand the field of view, the expansion has some limitations and APT system tracking is not easily achieved. In addition, the layout of nodal units and their adjustment become more difficult because multiple nodal units are involved; hence, the design needs further optimization and improvement.

Three-Concentric-Sphere Structure

The three-concentric-sphere structure, shown in Fig. 11.18, is designed by embedding three concentric-sphere structures into the lens to form a monolithic structure. Each concentric-sphere structure contains a lens and a reflective surface, and these surfaces work together to make a lens with both horizontal and pitch fields of view. By combining multiple such lens structures, multiple-target tracking and communication transmitting and receiving can be achieved. The system can reach a 120° field of view in both the horizontal and pitch directions, enabling it to be used in application scenarios with large field-of-view requirements, such as satellite communications.

By placing the unit-transceiver terminal on a tracking track, it is possible to transmit and receive communications to and from multiple targets. When the target

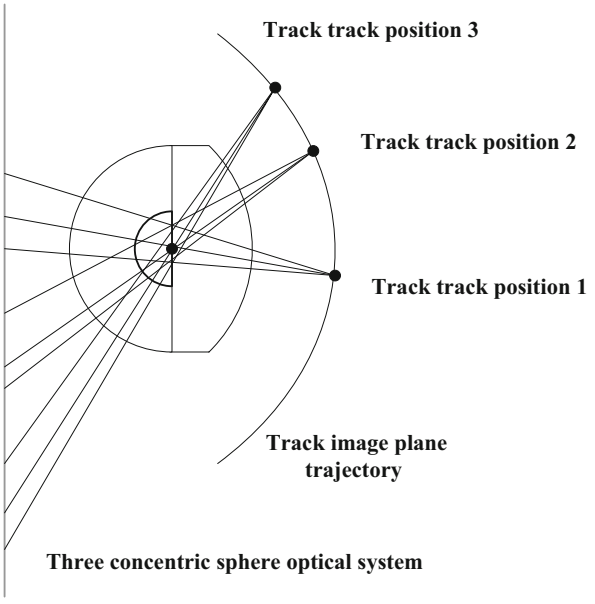


Fig. 11.18 Schematic diagram of the three-concentric-sphere structure [17]

moves into the jurisdiction of another subsystem, the system can quickly "take over" and handle the tracking and communication tasks, thus enabling a seamless communication connection [95]. It has a more efficient tracking and targeting capability than conventional optical antennas and identifies and tracks targets more quickly, thus making the work easier and more efficient.

This structure can be used to realize multi-point space laser communication in a defined orbit, which makes laser-communication networking simpler and more convenient. In addition, the optical antenna has a simple and lightweight structure, which is suitable for short-range laser-communication networks. During use, the structure can maintain good stability and high reliability and successfully work in complex environments [96].

Rotating-Paraboloid Structure

The rotating-paraboloid structure is shown in Fig. 11.19. It consists of a communication target, optical-relay system, detector, and transmitter. The rotating-paraboloid structure is obtained by using a reflective rotating paraboloid as the base surface type [97]. Its advantage is that it has a very wide horizontal communication range. The communication rays are directed to the focal point of the parabolic surface from different directions and are incident on the rotating-paraboloid structure. After the action of the rotating paraboloid, the light rays are emitted in a direction parallel to its axis and then incident onto the optical-relay system.

This reflective rotating paraboloid can achieve a horizontal communication range of 360° and a pitch communication range of 100°. Owing to the rotating-paraboloid structure, the antenna can be flexibly rotated and adjusted to adapt to different communication scenarios and is widely used in radar, aircraft, satellite receiving stations, and other fields. The rotating-paraboloid structure, with a compact design, wide communication range, and high communication efficiency, provides stable communication transmissions. This makes multi-target communication easy and efficient, enabling simultaneous one-to-many laser communications [86].

Development Trends

One-to-many wireless laser communication is a high-speed, high-quality communication method that has become a key technology for many applications. Its future development trend mainly includes the following points:

1. Networking: As the performance of optical wireless communication systems continues to gradually improve, from the current point-to-point mode to point-to-multipoint, building optical wireless communication networks is a direction of rapid development.

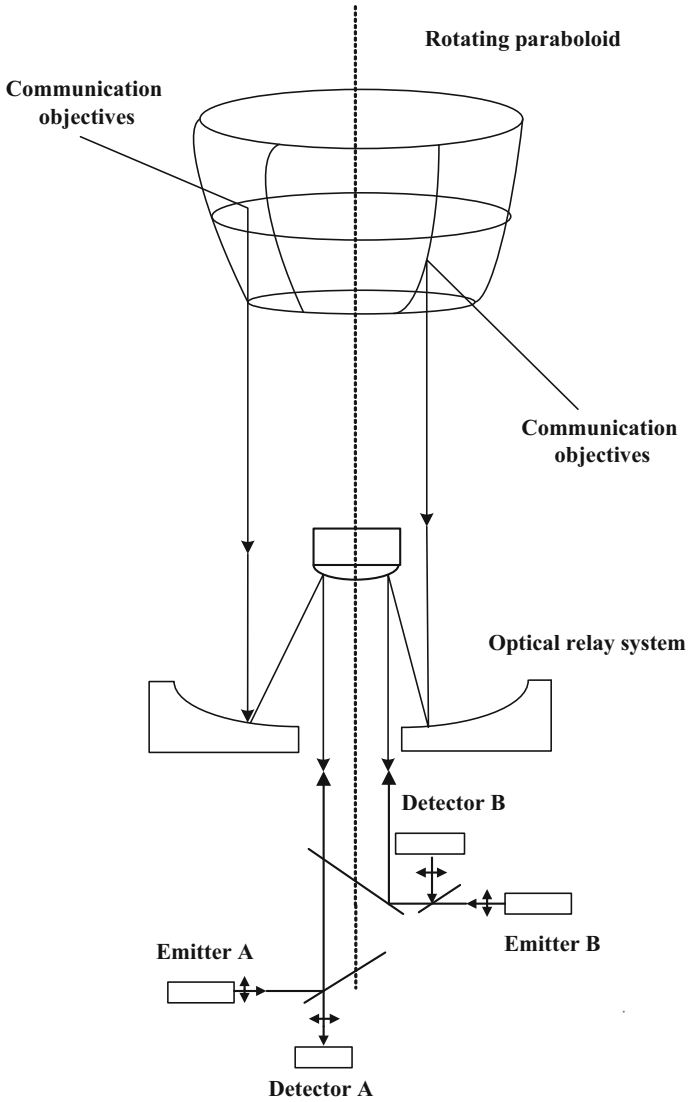


Fig. 11.19 Schematic diagram of the rotating-paraboloid structure [97]

2. High speed: With the development of technology, the transmission rate of one-to-many laser communication will gradually increase to achieve more efficient transmissions.
3. Security: To ensure the security of communications, the future will introduce more secure data-encryption technology to strengthen security measures.
4. Applicability: In the future, one-to-many laser-communication technology will be applied in a wider range of fields, such as military, medical, intelligent

transportation, etc., to provide communication services with a high transmission rate, high reliability, and high stability.

Generally, one-to-many laser-communication technology will become popular and become one of the main communication means in various fields, while more efficient and secure technologies and applications will be continuously introduced.

Summary

Space laser-communication networking requires capabilities, such as one-to-multi-point communication, dynamic communication, and duplex communication. In this chapter, the basic characteristics of different types of optical antennas were introduced and the current status of one-to-many laser communication at home and abroad and its development were discussed.

Moreover, several different types of structures were presented, namely, the multi-optical terminal combination type, multi-reflector superposition type, TSAT networking, dual-optical wedge group combination type, fiber-array structure, multi-transmitting and multi-receiving structure, three-concentric-sphere structure, and rotating-paraboloid structure. The features are compared in Table 11.5.

One-to-many communication technology has become an important research direction in the field of optical wireless communication, with broad application prospects. The future research directions of optical wireless communication systems with one-to-many optical antennas are likely to take the following paths:

1. Research on the tracking-control technology of one-to-many wireless laser communication should be strengthened.

Table 11.5 Comparison of the characteristics of each structure

Designation	Strengths and weaknesses
Multi-optical-terminal combination type	Fast capture and high-precision tracking; high weight
Multi-reflector superposition type	Low energy efficiency; not widely available
TSAT networking	High-bandwidth information communication capability
Double optical-wedge group combined type	Large field of view, dynamic tracking
Fiber-array structure	Large focal plane; small field of view, serious stray light, more difficult to arrange transceiver devices
Multi-transmit, multi-receive structure	Field of view is still limited and APT system tracking is more difficult to achieve
Three-concentric-sphere structure	Simple structure, light weight, and miniaturization
Rotating-paraboloid structure	Wide communication range, simple structure, long communication distance

2. One-to-many optical wireless communication should move toward greater coverage, higher transmission rates, and wider application prospects, and multipoint-to-multipoint communication, based on one-to-many communication, should be investigated.
3. The one-to-many optical wireless communication system optical-antenna design is a key technology to achieve one-to-many communication. This technology provides the feasibility and possibility for establishing optical wireless communication grouping networks.
4. Building an integrated global information network in space and air is an inevitable development trend, and wireless laser-communication networking technology is the most effective means to achieve this goal.
5. Traditional wireless laser communication can achieve point-to-point communication; however, it is difficult to meet the communication needs of multiple devices, thus limiting the development of optical wireless communication networking. One-to-many optical wireless communication technology can effectively solve the problems that point-to-point communication cannot solve.

References

1. Zhang Xuebin. Design of optical receiving antenna for visible light wireless communication [D]. Beijing: School of Optoelectronics and Beijing University of Technology, 2018.
2. Hu Yao. Research on the design of optical antenna for Optical Wireless Communication system [D]. Wuhan: Huazhong University of Science and Technology, 2018.
3. Sun Wenchang, Yuan Dan. Research and progress of wide-view optical receiving antenna [J]. Popular Science and Technology, 2012, 14(06): 54–58.
4. Cai Zheng, Zhang Chao, Fan Fan. The history and prospect of astronomical telescopes–and on the Tsinghua Wide Field Survey Telescope (MUST) [J]. Experimental Technology and Management, 2021, 38(05): 1–9+38.
5. Xue Qingsheng. Optical design of a satellite-borne wide-field differential absorption imaging spectrometer [J]. Journal of Optics, 2015, 35(01): 338–345.
6. Jiang Xiwen. Research on cooled off-axis reflective optical system [D]. Changchun: Changchun University of Technology, 2016.
7. Qian Liqiong. Design of large-aperture optical antenna [D]. Chengdu: University of Electronic Science and Technology, 2019.
8. Zhang Y, Ai Y. Drawback of Cassegrain antenna system based on space optical communication [J]. Infrared and Laser Engineering, 2005, 34(5): 560–563.
9. Li YJ J, Zhang Y. Co-aperture optical system for imaging and laser transmitting [J]. Chinese Optics, 2015, 12(02): 220–226.
10. Bing J, Guohua C, Qiongying L, et al. Optical design of tracking/guiding system with multi-spectrum and common aperture [J]. Infrared and Laser Engineering, 2017, 46(2): 218001–0218001 (7).
11. Xi Xiaodong, Wu Hanping. Design of near-surface UV communication transmitting optical system based on Cassegrain structure [J]. Applied Optics, 2017, 38(02): 205–2010.
12. Li Hongzhang. Study of laser radiation detection optics in optical countermeasure systems [J]. Optical countermeasures and passive interference, 1996, (04): 21–51.
13. Zhang Ruirui, Shen Weimin. Ring aperture folding type ultra-thin lens [J]. Infrared and Laser Engineering, 2012, 41(05): 1306–1311.
14. Guo Yongxiang, Li Yingcai, Liang Tianmei, et al. A large field-of-view off-axis triple reflection optical system [J]. Journal of Optics, 2010, 30(09): 2680–2683.

15. Juranek H J, Sand R, Schweizer J, et al. Off-axis telescopes: the future generation of Earth observation telescopes [J]. SPIE, 1998, 3439: 104–118.
16. Zhao Wencai. Design of an improved off-axis triple inverse optical system [J]. Optical Precision Engineering, 2011, 19(12): 2837–2843.
17. Liang Jingyuan, Kang Weilong, Ke Xizheng, et al. Research progress of optical antenna technology for free-space optical communication systems [J]. Optical Communication Technology, 2022, 46(04): 1–11.
18. Li Guang, Wang Jianye, Shao Hui, et al. Design of continuous zoom atmospheric coherence length measurement system [J]. Applied Optics, 2012, 33(03): 474–4710.
19. Zhang Lei. Research and implementation of star-guiding algorithm based on astronomical camera and equatorial instrument [D]. Hangzhou: Zhejiang University, 2018.
20. Gao Zhan, Zhang Yibo. Optical antenna system in free-space optical communication system [J]. Modern Telecommunication Technology, 2003, (03): 24–28.
21. Fan Lina, Zhu Aimin, Liu L, et al. Optical design of laser beam expanding telescope [J]. Infrared, 2007, (08): 20–22.
22. Xiong Yanjian, Wu Hanping, Lv Zhaosun, et al. Military infrared optical system performance and its structure form technical analysis [J]. Infrared technology, 2010, 32(12): 688–695+708.
23. Yang H, Hu Y, Li C, et al. Optimum design for optical antenna of space laser communication system [A]. IEEE. 2006 International Conference on Communications, Circuits and Systems [C].US: IEEE, 2006:2016-20110.
24. Ma Zixuan, Li Xuyang, Ren Zhiguang, et al. Design of ultra-compact detection optical system with large field of view [J]. Optical Precision Engineering, 2020, 28(12): 2581–2588.
25. Luo Hui, Li Jie, Li Jincheng, et al. Research on near-surface ultraviolet communication transceiver/transmitter integrated optical system technology [J]. Optical Technology Applications, 2021, 36(06): 10–23.
26. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress on adaptive optics for Optical Wireless Communication systems at Xi'an University of Technology [J]. Intense Laser and Particle Beam, 2021, 33(8): 30–52.
27. Clarke S, Prenger R, Ross M. Experimental results with a prototype three-channel multiaccess transceiver lasercom terminal [A]. SPIE. Free-Space Laser Communication Technologies [C]. Los Angoles: SPIE, 1993:128–138.
28. Zhizhong Y, Jingyi C, Xianwei Z. Development and technology of U.S. military satellite communications systems [J]. Communications Technology, 2009, 41(11): 55–59.
29. Sherman M P, John A T. Point-to-Multipoint Free-Space Optical Wireless Communication System. US6445496B1 [P]. 20029-3.
30. Li Z, Wen C. Research on telescope array for inter-satellites laser communication [J]. Proceedings of SPIE- The International Society for Optical Engineering, 2006, 6351(6): 58–66.
31. Zhang Cheng, Gu Wenbo, Liu Di, et al. Preliminary study on multi-point terrestrial reception scheme for satellite-ground optical communication [J]. Journal of Photonics, 2008, 16(02): 275–279.
32. Zhang W L, Pan W, Luo B, et al. One-to-many and many-to-one optical chaos communications using semiconductor lasers [J]. IEEE Photonics Technology Letters, 2008, 20(9): 712–716.
33. Li Ke, Wang Jiangan, Yao Yao. Communication characteristics of laser sea surface diffuse reflection grouping under Kirchhoff approximation [J]. Intense Lasers and Particle Beams, 2011, 23(07): 1775–17710.
34. Jiang Huilin, Hu Yuan, Song Yansong, et al. Research on space laser communication networking optical terminal technology [J]. Space Return and Remote Sensing, 2011, 32(5): 52–510.
35. Jiang Huilin, Hu Yuan, Ding Ying, et al. Research on the optical principle of space laser communication networking [J]. Journal of Optics, 2012, 32(10): 56–60.
36. Jing Y H, Jiang H, Hu Y, et al. Space Laser Communication Network [J]. International Conference on Space Optical Systems and Applications, 2012, 9(12):4–6.
37. Sun Xiaodong. Research on the follow-up mechanism of one-to-many laser communication terminals [D]. Changchun: Changchun University of Technology, 2014.

38. Xiao Pan. All-optical capture technology for one-to-many laser communication terminators [D]. Changchun: Changchun University of Technology, 2014.
39. Yu Xiaonan, Tong Shoufeng, Jiang Huilin, et al. Research on laser communication network control system [A]. Chinese Society for Space Science. Proceedings of the 2013 Symposium on Space Optics and Mechatronics of the Chinese Space Science Society [C]. Chinese Society for Space Science: Chinese Society for Space Science, 2013: 386–391.
40. Yu Xiaonan, Tong Shoufeng, Dong Yan, et al. Single-beam tracking subsystem for space laser communication networking [J]. Optical precision engineering, 2014, 22(12): 3348–3353.
41. Jiang Huilin, Liu Xianzhu, HU Yuan, et al. Reflections on several key issues of the integrated information network between heaven and earth [J]. Journal of Military Engineering, 2014, 35 (S1): 96–100.
42. Meng L X, Li L, Zhang L, et al. Research on optic antenna of space laser communication networking [A]. SPIE. International Symposium on Photoelectronic Detection and Imaging 2013 [C]. Laser Communication Technologies and Systems: SPIE, 2013, 8906: 589–596.
43. Meng Lixin, Zhao Dingxuan, Zhang Lizhong, et al. Design of heel catching mechanism for laser communication networking [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2014, 37(06): 70–76.
44. Meng Lixin. Research on capture and tracking technology in airborne laser communication [D]. Changchun: Jilin University, 2016.
45. Zheng Yang, Fu Yuegang, Hu Yuan, et al. Study on dynamic alignment of four mirrors based on space laser communication networking [J]. China Laser, 2014, 41(1): 103–1010.
46. Zhao Kehan. One-to-many laser communication optical antenna tracking performance test method research [D]. Changchun: Changchun University of Technology, 2018.
47. Zhang Lingjun, Liu Yunqing, Dong Yan. Research on the application of Kalman filtering in spatial multipoint optical communication tracking system [J]. Science and Technology Information, 2014, 358(01): 14–18.
48. Zhang Lingjun. Research on the key technology of space laser communication network tracking control [D]. Changchun: Changchun University of Technology, 2016.
49. Zhao Kehan, Zhang Lizhong, Wang Xia. Research on optical antenna tracking performance testing method [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2015, 38(01): 57–60.
50. Yu Xiaonan, Tong Shoufeng, Jiang Huilin, et al. Servo system of optical relay antenna for multipoint laser communication [J]. Journal of Optics, 2015, 35(09): 176–183.
51. Jiang Huilin, Jiang Lun, Song Yansong, et al. Research on optical follow-up technology for one-to-multipoint simultaneous space laser communication [J]. China Laser, 2015, 42(04): 150–159.
52. Jiang Lun, Li Na, Fu Qiang, et al. Application of one-to-many laser communication technology in formation flight constellation [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2016, 39(06): 5–8+18.
53. Zhang Tao, Fu Qiang, Li Yahong, et al. Design of confocal reflective optical antenna for multiplatform laser communication networking [J]. Journal of Optics, 2015, 35(10): 220–226.
54. Ke Xizheng, Lei Sichen, Wang Tianyu. Fiber optic array-based Cassegrain transceiver integrated antenna [J]. Advances in Laser and Optoelectronics, 2015, 52(12): 49–55.
55. Zhang T. Multi-platform laser communication networking optical antenna system design [A]. SPIE. Optical Measurement Technology and Instrumentation [C]. Beijing: SPIE, 2016: 274–2710.
56. Zhang Tao, Fu Qiang, Jiang Huilin, et al. Effect of energy transceiver efficiency on the performance of confocal reflective optical communication grouping antenna system [J]. Journal of Optics, 2016, 36(12): 42–50.
57. Ma Guodong, Zhang Lizhong., Zhang Xueyao. Joint simulation of one-to-many laser tracking antennas based on ADAMS and Matlab [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2016, 39(03): 78–82.
58. Jiang Lun, Hu Yuan, Wang Chao, et al. Research on one-point-to-multipoint simultaneous space laser communication optical system [J]. Journal of Optics, 2016, 36(05): 37–43.

59. Zhang T, Mao S, Fu Q, et al. Networking optical antenna of space laser communication [J]. Journal of laser applications, 2017, 29(1): 012013–012123.
60. Zhao Y, Xu J, Wang A, et al. Demonstration of data-carrying orbital angular momentum-based underwater wireless optical multicasting link [J]. Optics Express, 2017, 25(23): 28743–28751.
61. Li Xiang, Wang Shouda, Zhang Jiaqi, et al. Carrier laser communication optical antenna [J]. Infrared and Laser Engineering, 2019, 48(11): 221–229.
62. Lou Nan. Structural design of one-to-many wireless laser communication abutments [D]. Changchun: Changchun University of Technology, 20110.
63. Cao Hongyun. Research on coarse tracking technology for satellite-based one-to-many laser communication optical transceivers [D]. Changchun: Changchun University of Technology, 2020.
64. Wang Shouda. Research on optical abutment technology for multi-point simultaneous laser communication terminal [D]. Changchun: Changchun University of Technology, 2020.
65. Song G, Wang T, Dong F, et al. Transmission characteristics of 24.5 Gb/s atmospheric laser communication based on optical frequency comb [J]. Optics Communications, 2020, 465: 125602–125611.
66. Wang Junyao, Song Yansong, Tong S F, et al. Linkage tracking control technology of reflector for space laser communication networking [J]. China Optics, 2020, 13(03): 537–546.
67. Wang J Y, Song Y, Jiang H, et al. Research and Dynamic Demonstration Test of One-Point to Multi-Point Space Laser Communication System [A]. IEEE. 2021 13th International Conference on Advanced Information Technology [C]. Guilin: IEEE, 2021: 143–146.
68. Wang T, Luan Y, Zhao X, et al. Effect of Platform Vibration Characteristics and Atmospheric Turbulence Characteristics on Free-space Optical Communication Systems [A]. IEEE. 2021 International Conference on Electronic Information Engineering and Computer Science [C]. Guilin: IEEE, 2021: 39–42.
69. Bai Yangyang, Cen Yuanyao, Meng Lixin, et al. Research on control technology of space laser communication networking from optical transceivers [J]. Journal of Optics, 2021, 41(14): 18–25.
70. Fan Tao, Zhang Lizhong, Meng Lixin, et al. Research on fine tracking control technology of main optical transceiver for space optical communication networking [J]. Laser Journal, 2021, 42(11): 112–118.
71. Li Xiaoming, Wang Longming, Zhu Guoshuai. Optimization of support parameters for integrated SiC/Al pendulum mirror for laser communication [J]. Infrared and Laser Engineering, 2021, 50(11): 194–201.
72. Li Xiaoming, Wang Longming, Li Xiang, et al. Pendulum mirror surface shape optimization for laser communication antenna integration [J]. China Laser, 2021, 48(01): 142–150.
73. Zhao Baiqiu, Yu Xiaonan, Dong Yan, et al. Tracking performance of reflector for space laser communication networking [J]. Advances in Laser and Optoelectronics, 2021, 58(09): 203–212.
74. Xu Miao, Shi Haodong, Wang Chao, et al. Research on integrated technology of multi-dimensional detection of space targets and laser communication [J]. China Laser, 2021, 48(12): 229–241.
75. Zhang Kaiqiang, Zhang Lizhong, Bai Yangyang, et al. Model predictive control of coarse tracking system for spatially grouped optical communication [J]. Laser Journal, 2022, 43(05): 17–22.
76. Ding C, Li C, Wang Z, et al. Free Space Optical Communication Networking Technology Based on a Laser Relay Station [J]. Applied Sciences, 2023, 13(4): 2567–2574.
77. Hu Yibo, Meng Lixin, Bai Yangyang, et al. Coarse tracking equivalent composite control technique for space laser communication [J/OL]. (2023-02-10) [2023-04-26]. http://kns.cnki.net/kcms/detail/31.1690.TN.202302010.1400.005.html.
78. Wang Lihui, Zhang Lizhong, Meng Lixin, et al. One-point-to-multipoint simultaneous laser communication with view-axis attitude solving method [J/OL]. (2023-03-09)[2023-04-20]. http://kns.cnki.net/kcms/detail/12.4494.tn.20230308.0929.002.html.

79. Jin Guang, Li Yanjie, Zhong Xing, et al. Design of common aperture optical system for space imaging and laser communication [J]. Optical Precision Engineering, 2014, 22(8): 8–16.
80. Dong Lu, Wang Jun. Space laser communication key technology and industrialization development [A]. Satellite Communication Committee of the Chinese Society of Communication, Satellite Application Committee of the Chinese Society of Astronautics. Proceedings of the 17th Annual Academic Conference on Satellite Communication [C]. Beijing, China: Chinese Society of Communication, 2021:359–365.
81. Zhang Yu, Gang Jianxun, Du Xi. Analysis of the current situation and development trend of space laser communication research [J]. Electronic test, 2015, (6): 93–94.
82. Ke Xizheng, Deng Lijun. Optical wireless communication [M]. Beijing: Science Press, 2016: 165–166.
83. Tang L Y, Li F J, Xie X L, et al. Study on irradiation-induced defects in GaAs/AlGaAs core–shell nanowires via photoluminescence technique [J]. Chinese Physics B, 2017, 26(08): 332–336.
84. Wang Heding. Development of electromagnetically driven fast reflector servo actuator [D]. Shanghai: Shanghai Jiao Tong University, 2019.
85. Peng Xujin, MA Jiaguang. Experiments and studies on the composite axis control system in optical precision tracking [J]. Optoelectronic Engineering, 1994, 21(5): 1–10.
86. Zhang Yalin, An Yan, Jiang Huilin, et al. Comparative study on the principle and method of point-to-multipoint optics for space laser communication [J]. Journal of Military Engineering, 2016, 37(01): 165–171.
87. Jiang Huilin, Hu Yuan, Song Yansong, et al. Research on space laser communication networking optical terminal technology [J]. Space Return and Remote Sensing, 2011, 32(05): 52–510.
88. Liu Tiefeng, Zhang Minghua. The U.S. Army transformational satellite communication system and its development [J]. Satellite and Network, 2011, (05): 62–63.
89. Zhao Jing, Zhao Shanghong, Li Yongjun, et al. Research progress of interplanetary laser link data relay technology [J]. Infrared and Laser Engineering, 2013, 42(11): 3103–3111.
90. Yin Zhizhong, Chen Jingyi, Zhou Xianwei. Development of the U.S. Army satellite communication system and its technical research [J]. Communication Technology, 2009, 42(11): 55–59.
91. Dai Zhengshuang, Jiang Lun, Song Zhihua, et al. Coherent detection mixing system with shaping elements [J]. Advances in Laser and Optoelectronics, 2021, 58(07): 132–139.
92. Presby H M, Tyson J A. Point-to-multipoint free-space Optical Wireless Communication system: US6445496B1[P]. 2002-09-03.
93. Wang Ling, Chen Xi, Dong Feng. Development level and development trend of space laser communication optical transceiver [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2016, 39(02): 39–45.
94. Sparrold S W, Upton E L, Okorogu A O. Free space point-to-multipoint optical communication system and apparatus: US6912360B1[P]. 2005-06-29.
95. An Yan, Zhang Yalin, Zhao Yiwu, et al. Structural analysis and optical design of three-concentric spherical antenna network [A]. Department of Information and Electronic Engineering, Chinese Academy of Engineering. Proceedings of the National Defense Optoelectronics Forum Workshop on Spatial Optical Information and Networking Technology and Applications [C]. Changchun, Jilin, China: Chinese Society of Optical Engineering, 2015: 4–11.
96. An Yan, Li Xinhang, Zhao Yiwu, et al. Analysis of follow-up error of three-concentric sphere optical system [J]. China Optics, 2016, 9(06): 687–694.
97. Zhang Yalin, An Yan, Wang Chao, et al. Study of rotating parabolic substrate surface type in space laser communication networking [J]. Journal of Optics, 2015, 35(07): 94–99.

Acquisition, Pointing, and Tracking

12

Contents

Introduction 358
Foreign Research Progress 360
Domestic Research Progress 361
Research Progress in the Automatic Aiming of Optical Wireless Communication Systems at Xi'an University of Technology 364
Coarse- and Fine-Grading Control System 367
Beam Detection System 369
Automatic Tracking Control System for Space-Based Relay Optical Wireless Communication 369
Field Experiments for Acquisition, Aiming, and Tracking in Optical Wireless Communication 372
Non-common Visual Axis Two-Dimensional Mirror Rapid Aiming System 375
Chapter Review 386
Summary 388
References 389

Abstract

Optical wireless communication refers to the technology of transmitting information in free space using light waves as a carrier, which has the advantages of high bandwidth, low cost, and high security. The acquisition, pointing, and tracking (APT) system is the premise of establishing an optical wireless communication system. A simple, reliable, and dynamic APT system can overcome the impact of mechanical platform vibration and external environment changes on the optical wireless communication system. Therefore, it is necessary to conduct in-depth theoretical and experimental research on the APT system, so as to design a capture, aiming, and tracking method suitable for optical wireless communication. This chapter analyzes the domestic and foreign research achievements in capturing, aiming, and tracking, and introduces the work done by Xi'an University of Technology in the field of automatic aiming. It mainly includes the progress of initial acquisition system, non-common visual axis control system,

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_12

beam detection system, etc. At the same time, the field experiments of 1.3 km, 5.2 km, 10.2 km, and 100 km distance links are introduced to verify the effectiveness of the APT system. Finally, the development of APT in optical wireless communication is prospected.

Keywords

Optical wireless communication · Acquisition pointing and tracking · Beam detection · Non-common visual axis control

Introduction

Optical wireless communication uses light waves as carriers to transmit information in free space. It offers the advantages of high bandwidth, low cost, and high security. An acquisition, pointing, and tracking (APT) system is the basis for establishing an optical wireless communication system. An APT system with simple, reliable, and good dynamic performance can overcome the impact of mechanical-platform vibrations and external environment changes on optical wireless communication systems. Therefore, it is necessary to conduct more in-depth theoretical and experimental research on APT systems to design an acquisition, pointing, and tracking method suitable for optical wireless communication.

This chapter analyzes the research results on acquisition, pointing, and tracking at home and abroad and introduces the work of Xi'an University of Technology in automatic aiming. It includes the progress of the initial acquisition system, non-common axis-of-sight control system, beam detection system, etc. In addition, outfield experiments of 1.3 km, 5.2 km, 10.2 km, and 100 km distance links are introduced, which verify the effectiveness of the APT system. Finally, the development of APT in optical wireless communication is projected.

Optical wireless communication has the advantages of both microwave and optical fiber communications [1]. It is an important means to realize space-based, air-based, and land-based communication links [2] and has broad development prospects in military-security communication, emergency-rescue communication, and other fields.

However, owing to the influence of mechanical jitter and various uncertain factors in the atmospheric channel (such as atmospheric turbulence, rain, snow, and dust), the spot position of the receiver may be offset or even lost. This increases the bit error rate of the system and ultimately leads to a decline in performance, causing difficulties in establishing and maintaining optical wireless communication links [3]. Therefore, it is necessary to establish a set of acquisition, pointing, and tracking (APT) systems to prevent communication link interruptions.

In an optical wireless communication system, the optical elements on the transmitter and receiver platforms are required to share the same axis of view in real time. This process is typically referred to as automatic aiming [4]. A structural diagram of

the optical wireless APT system is shown in Fig. 12.1. The APT system includes coarse and fine tracking and aiming structures and servo units. After the initial pointing of the transmitter and receiver is completed, the signal and beacon lights pass through the beam-combining mirror, are collimated, and are output through the transmitting antenna.

The acquisition device at the receiving end starts to scan the synthetic beam to complete the acquisition work. The beam splitter at the receiving end guides the beams through the receiving end to the tracking and communication detectors. When the light beam shines on the photosensitive surface of the detector at the receiving end, the sight axes at both ends of the transmitter and receiver are aimed by the drive of the servo mechanism.

Finally, according to the real-time feedback spot information at the receiving end, the coarse tracking structure adjusts the direction of the beam to make it perpendicular to the center of the detector. The fine tracking structure further eliminates the tracking residual that was not eliminated by the coarse tracking structure, to ensure the real-time collimation of the line of sight (LOS) at both ends of the transmitter and receiver.

This chapter introduces the research progress and course of APT in optical wireless communication systems at home and abroad, summarizes the research achievements of Xi'an University of Technology in this field, and looks forward to the application prospects and research directions of APT systems in optical wireless communication, according to existing research theories.

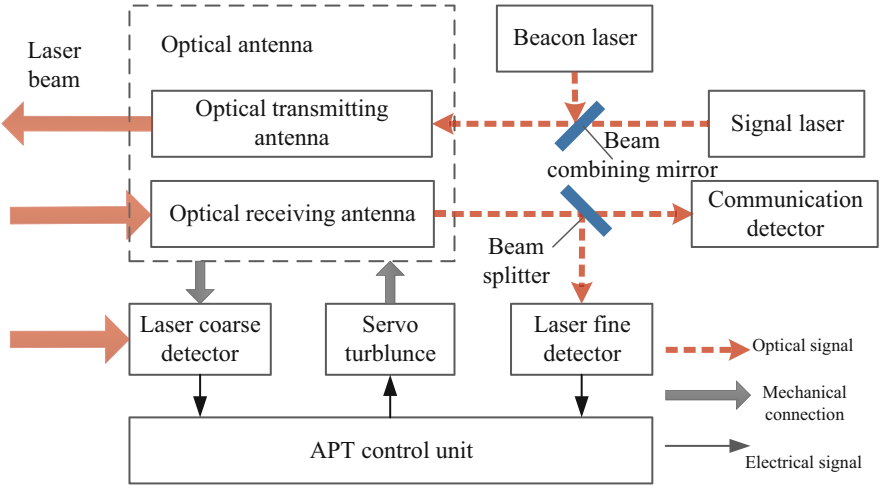


Fig. 12.1 Schematic diagram of an optical wireless communication APT system structure [5]

Foreign Research Progress

An APT system is an indispensable part of an optical wireless communication system. Its role is to ensure that the beams of the transmitter and receiver are fully aligned in real time, to establish a stable optical wireless communication link. The development of APT systems has been almost synchronous with the development of optical wireless communication systems.

In the 1970s, the Jet Propulsion Laboratory (JPL) in the United States conducted experimental research on space-based optical wireless communication systems. The Deutsches Zentrum für Luft- und Raumfahrt (DLR), European Space Agency (ESA), National Space Development Agency of Japan (NASDA), and other institutions successively carried out satellite–ground experiments on air–ground optical communication links.

In 1985, NASDA developed the laser utilizing communication equipment (LUCE) system, in which the detection system used for alignment and tracking consisted of a charge-coupled device (CCD) and a four-quadrant detector. The experiment showed that the aiming and tracking accuracies of the system were better than 1 mrad [6] in a strong vibration environment.

In 1994, the JPL developed the optical communication demonstrator (OCD) system that realized fast, accurate aiming and stable beam tracking with a data-transmission rate of 250 Mbit/s [7].

In 1994, the Ministry of Posts and Telecommunications (MPT) of Japan developed laser communication equipment (LCE). The coarse tracking structure of the system adopted area array CCD detection, and the coarse tracking accuracy was 32 μrad. The precision tracking structure used a four-quadrant detector to measure the position information of the beam; its accuracy reached 2 μrad, and it had a jitter compensation capability of 200 Hz [8].

In 1999, Biswas designed a satellite laser communication terminal (LCT) that used an independent transmitting and receiving antenna. The capture sensor used an area array CCD, and a servo turntable for aiming and tracking. The frame rate of the CCD operation in tracking mode reached 1.6 kHz [9].

In 2002, the ESA established an inter-satellite optical communication link using the semiconductor laser inter-satellite link experiment (SILEX) system. The system used a coarse and fine composite axis alignment system composed of a direct current (DC) stepper motor and a fast mirror. Its tracking accuracy reached 2 μrad [10].

In 2004, Guelman adopted a two-level hierarchical system control, with the lower layer controlling the pan, tilt, and zoom (PTZ) of the electro-optical transceiver. The upper layer was a fast closed-loop control system to control the beam width and direction, and realized laser communications between an unmanned aerial vehicle (UAV) and the ground [11].

In 2004, the Lincoln Laboratory of the Massachusetts Institute of Technology and JPL jointly developed the Mars laser communications demonstration system (MLCD). The system adopted a photon-counting detector and telescope receiving-array structure. A combined aiming and tracking system was adopted to improve the aiming accuracy, and the system communication rate reached up to 10 Mbit/s [12].

In 2010, the DLR designed a beacon-free laser communication terminal (LCT). In the process of aiming and tracking, a beacon light was not used, and a servo turntable control mirror with only two dimensions was used for spiral scanning and signal light acquisition. This reduced the complexity of the airborne laser communication terminal. The average tracking error was 226 μrad [13].

In 2012, Schmidt et al. designed a small laser communication terminal consisting of an objective lens providing a wide field of view and a driver for controlling the beam-steering mirror. The system provided a wide field of view of 46°, based on fast aiming [14].

In 2013, the DLR conducted an air-to-ground optical wireless communication experiment. The airborne laser terminal included a coarse aiming structure and fine aiming mechanism. Experiments showed that the system could reduce the influence of turbulence and microvibration on the tracking accuracy to 200 μrad [15]. In the same year, the DLR conducted an airborne laser communication experiment between the "Wind" fighter and a ground terminal. The communication link distance was 50 km, the data-transmission rate was 1.25 Gbit/s, and the APT system terminal carried by the aircraft achieved stable tracking at 79 km [16].

In 2016, Quintana et al. designed a tracking system consisting of a coarse tracking module, composed of a CCD camera, and a fine-tracking module, composed of a spatial light modulator and a position-sensitive detector (PSD), and applied the system to an air–ground laser communication experiment. A 2-Mbit/s optical communication link was established within 300 m, with a bit error rate of approximately 2×10^{-4} [17].

In 2020, Antonello et al. designed a high-precision aiming and tracking system for in-orbit laser communications. The system adopted coarse and fine aiming structures with a weight of 1.8 kg. By simulating an external interference environment, the aiming error of the system was found to be less than 10 μrad [18].

A summary of the foreign research progress is presented in Table 12.1. As research on optical wireless communication APT systems has been conducted, their application scenarios have gradually developed from ground stations to airborne and inter-satellite optical communication platforms. The accuracy requirements of the tracking and pointing systems have also increased.

Domestic Research Progress

Since the 1990s, many scientific research institutions and universities in China have conducted in-depth research in the field of optical wireless communication and have made certain progress in rapid aiming systems, composite axis control systems, and airborne laser communication systems. At present, the main research institutions in China include the Changchun University of Technology; Harbin Institute of Technology; Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences; University of Electronic Science and Technology; Wuhan University; and Xi'an University of Technology.

Table 12.1 Foreign research progress

Year	Researcher/ organization	Research progress	Benefits/parameters
1985	NASDA	LUCE system	Tracking accuracy better than 1 mrad
1994	JPL	OCD system	250-Mbit/s communication rate
1994	MPT	LCE system	Tracking accuracy better than 32 μrad and 2 μrad
1999	Biswas	LCT system	1.6-kHz CCD working frame rate
2002	ESA	Composite aiming system applied to SILEX	Tracking accuracy was 2 μrad
2004	Guelman	Composite APT system	Composite shaft system adopted for the first time
2004	MIT NASA	OLCD system	10-Mbit/s communication rate
2010	DLR	Laser communication terminal	226-μrad average tracking error
2012	Schmidt	Small laser communication terminal	Capture field of view was 46°
2013	DLR	Experiment on ground-to-air laser communication	Distance: 79 km; transmission rate: 1.25 Gbit/s
2016	Quintana	Airborne laser communication system	Communication rate: 2 Mbit/s
2020	Antonello	Satellite laser communication system	Aiming error less than 10 μrad

In 1999, Liu et al. designed an automatic aiming scheme for optical wireless communication systems. In this scheme, a reflector was used as the beam stabilizer to eliminate the optical axis deviation at both ends of the transmitter and receiver, caused by beam drift and mechanical vibrations in the system [4].

In 2000, Jiang et al. designed a prototype of a wireless laser communication system. The transmitter used a semiconductor laser with an 855-nm wavelength. The maximum communication rate of the system was 2048 Kbit/s [19].

In 2002, Hu et al. designed a free-space optical communication ground demonstration system. The system used a tracking mechanism composed of a DC torque motor to achieve stable high-precision and broadband operation. The field experiment results showed that the system acquisition time was 180 s, and the tracking and pointing accuracies in the azimuthal and pitch directions were 38.6 μrad and 40.4 μrad, respectively [20].

In 2005, Ai et al. designed a space light APT scheme that used a CCD sensor as an image processor and a digital signal processor to control and drive the composite axis system to realize tracking. The experimental results indicated that the relative error of the angle measurement was approximately 1.3% [21].

In 2007, Jiang et al. optimized the bandwidth of a composite axis coarse tracking servo system and simulated an interstellar optical wireless communication experiment. When the servo bandwidth was 16 Hz, the coarse tracking and fine tracking errors were 60 μrad and 4 μrad, respectively [22].

In 2008, Zhang et al. designed an APT experimental device for deep-space noncooperative targets and analyzed the motion characteristics of the system's tracking structure, and the factors affecting the acquisition and tracking time after the satellite deviated from the center of the spot. The results showed that during the acquisition and tracking process, the angular velocity of the motor met the change rule of acceleration before deceleration. The angular acceleration of the motor and the detection accuracy of the detector directly affected the acquisition and tracking time of the system [23].

In 2008, Pan et al. designed an automatic pointing system for uplinks and downlinks. By adjusting the angle of the tilt mirror, the transmitter was aimed at a CCD goniometer, and the aiming accuracy of the optical axis at both ends of the transmitter and receiver reached 20.52 μrad [24].

In 2011, Ma et al. realized optical communication between the "Ocean 2" satellite and the ground. The experimental results showed that the system uplink and downlink single-channel data rates were 2 Mbit/s and 504 Mbit/s, respectively [25].

In 2011, Song et al. analyzed the impact of external constraints on an airborne laser communication system and conducted an optical wireless communication experiment between two aircraft. The aircraft flew at a height of 700 m, with a communication rate of 1.5 Gbit/s [26].

In 2012, Ai et al. designed a precise tracking system and conducted a 2-km laser communication experiment. The system was primarily composed of a complementary metal-oxide semiconductor (CMOS) camera, fast mirror, and host computer, which were used to compensate for the coarse tracking residual caused by atmospheric turbulence and platform vibration. The experimental results showed that the standard deviation of the spot offset could be controlled to 6 μrad [27].

In 2013, Qian et al. proposed a design scheme for a new type of spot detection camera that was based on a detector with a higher pixel duty ratio and reduced the impact of noise on the detector-positioning accuracy. The experimental results showed that the positioning error of noise in the system could be reduced to less than 0.007 pixels using this scheme, which met the system requirements [28].

In 2015, Meng et al. designed a two-stage (i.e., coarse and fine) composite tracking mechanism for airborne laser communication systems that could eliminate platform vibration. The simulation results showed that the coarse and fine tracking accuracies of the system were better than 23.97 μrad and 8.03 μrad, respectively [29].

In 2017, Zhang et al. designed a composite tracking system for airborne laser communication experiments using a single-loop proportional–integral–derivative (PID) algorithm to control the fast reflector. Experimental results showed that the tracking accuracy of the system was 10 μrad [30].

In 2017, Xu et al. proposed a fast-tracking method using a liquid-crystal optical phased array. A PID algorithm was introduced to control the actuator to deflect the incident beam and point it to the center of the detector for fast tracking. The experimental results showed that the system tracking accuracy under a stable tracking state was less than 13.6 μrad [31].

In 2018, Yu et al. studied the hysteresis characteristics of a piezoelectric tilt mirror in a space laser communication APT system. To overcome the influence of hysteresis on the positioning accuracy of the precision aiming system, a mathematical model of the improved Prandtl–Ishlinskii play hysteresis operator was proposed, and the system parameters were identified using the least-squares method. The experimental results showed that the tracking error of the system using this method was reduced by 80% [32].

In 2019, Cai et al. conducted a study on the coarse and fine pointing of a single-detection composite axis system and enabled the coarse and fine pointing of the transmitter and receiver by establishing a decoupling model. Experimental results showed that the tracking accuracy of the system using this method reached 10.69 μrad [33].

In 2019, An et al. from the 34th Research Institute of the China Electronics Technology Group Corporation designed an airborne laser communication system. The system adopted a combination of coarse and fine tracking methods and conducted an airborne laser communication experiment with a communication distance of 6.7 km and a communication rate of 1.25 Gbit/s through speed and position closed-loop control [34].

In 2020, Ren et al. studied the application of a four-quadrant detector in beam acquisition and tracking systems and achieved stable tracking by calculating the spot position on the four-quadrant detector. Experimental results showed that the tracking accuracy of the system using this scheme was better than 3 μrad [35].

In 2021, Li et al. designed a backpropagation (BP) neural network position detection algorithm and a multielement array position detection method to meet the performance requirements of coarse and fine tracking. The experimental results showed that the angular resolutions of the spot position detection system were 0.187 μrad and 0.903 μrad, respectively [36].

In 2022, Zhong proposed a new scanning acquisition method. The system used a fast mirror as an actuator, overcoming the problem of large overlapping areas in traditional scanning methods. The experimental results showed that the acquisition time of this method was approximately 35 s less than that of the traditional non-beacon scanning method, and the acquisition probability could reach 92.3% [37].

A summary of domestic research progress is shown in Table 12.2.

Research Progress in the Automatic Aiming of Optical Wireless Communication Systems at Xi'an University of Technology

To maintain the real-time collimation of an optical transceiver line of sight (LOS), it is necessary to design a fast and high-precision APT system [29]. A typical optical wireless communication APT system is shown in Fig. 12.2 [1]. The main controller completes the initial orientation according to the positions of the transmitter and receiver ends, and the image-processing unit calculates the deflection angle of the two-dimensional turntable according to the spot position information. The angle

Table 12.2 Domestic research progress

Year	People	Research progress	Benefits/parameters
1999	Liu and Shu	Automated pointing system	Stable effective bandwidth was 50 Hz
2005	Ke and Liu	Automated acquisition system	Established an APT system simulation model
2005	Ai and Zhou	Space optical APT system	Angle measurement error was about 1.3%
2007	Jiang and Tong	Bandwidth optimization design of a coarse tracking servo	Coarse and fine tracking accuracies were about 60 μrad and 4 μrad, respectively
2008	Pan and Zhang	Automated pointing system	Pointing accuracy was 20.52 μrad
2011	Ke and Hu	Beacon light spot detection system	Used PSD and CCD to design an APT system
2011	Song and Chang	Airborne laser communication experiment	Communication rate: 1.5 Gbit/s
2013	Qian	Spot detection camera	Impact of noise on positioning error was reduced to 0.007 pixel
2015	Meng and Zhao	Compound tracking system	Coarse and fine tracking accuracies were better than 23.97 μrad and 8.0 μrad
2016	Ke and Yang	Coaxial aiming detection method	Angle tracking accuracy was 34.6 μrad
2016	Ke and Zhao	Initial capture system	A position calibration point method was designed
2017	Zhang and Qiu	APT system applied to airborne laser communication	Tracking accuracy was 10 μrad
2019	Ke and Yan	Spot tracking system	Tracking accuracy was 5.4 μrad
2019	Cai and Kong	Compound axis aiming system	Tracking accuracy was 10.69 μrad
2020	Ke and Jing	Spot image detection algorithm	Realized beacon-free optical aiming in a 100-km experiment
2020	Ke and Zhang	Capture, aiming, and focusing system	Tracking accuracy of a 10.2-km experiment was 28.12 μrad
2020	Ren and Lu	Four-quadrant detector tracking system	Tracking accuracy was better than 3 μrad
2021	Ke and Yang	Two-dimensional mirror fast alignment system	Transmitter used a camera for calibration and completed aiming without returning control information
2021	Ke and Liang	Airborne laser automatic tracking system	Tracking accuracy was 5.4 μrad
2021	Li and Wu	Neural network/array position detection	Angular resolution of spot position detection system was 0.187 μrad/ 0.903 μrad

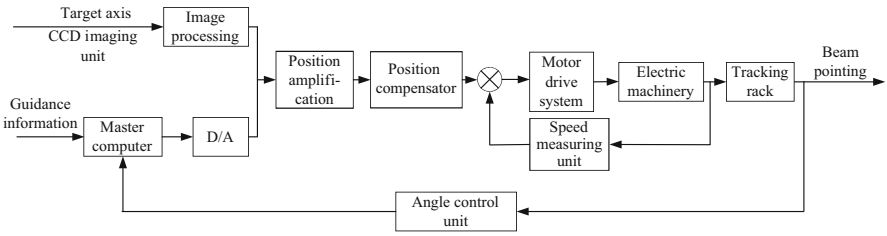


Fig. 12.2 Typical wireless laser communication APT system diagram [1]

control system feeds back the deflection angle information of the two-dimensional turntable in real time. The main controller then sends a control command to the motor to drive the two-dimensional turntable according to the feedback information to realize stable beam alignment and tracking.

Since 2000, Professor Xizheng Ke of Xi'an University of Technology has conducted in-depth theoretical research and experimental verification of typical optical wireless transmission technologies, such as optical wireless communication APT-system design, initial beam acquisition, precise spot aiming, and far-field spot image detection.

In 2005, Liu established and analyzed a simulation model for an APT system in atmospheric laser communication and designed an automatic beam acquisition system [30].

In 2011, Hu designed a beacon light spot detection scheme using a CCD [31].

In 2016, Yang proposed a coaxial aiming detection method, designed aiming control and tracking systems according to the method, and conducted field experiments [32]. In the same year, Zhao designed an initial acquisition system and conducted a 1.3-km outfield experiment [33]. In the same year, Xu designed a spot detection system and proposed a corresponding image-processing algorithm [34].

In 2017, Li proposed an optical axis aiming scheme that could effectively improve the detection and aiming accuracy of the system [35].

In 2019, Yan designed a spot tracking system and conducted a 5.2-km outfield tracking experiment. The experimental results showed that the tracking accuracy of the system could reach 5.4 μrad [36].

In 2020, Jing designed a spot image detection method and carried out a 100-km laser communication experiment [37]. In the same year, Zhang embedded a high-precision actuator in an APT system to achieve high-precision aiming and tracking, designed a focusing system, and carried out field experiments at 10.2 km and 100 km [38].

In 2021, Liang designed an APT system that could be carried on a UAV and carried out an airborne laser communication experiment by simulating an airborne experimental platform, with a tracking accuracy of 2.42 μrad [39]. In the same year, Ke et al. proposed a rapid aiming method. This method did not require the control signal to feed from the receiver back to the transmitter and could simultaneously establish uplinks and downlinks. Field experiments of 1.3 km and 10.2 km were conducted [40–41].

Coarse- and Fine-Grading Control System

A coarse- and fine-grading control system is composed of a two-dimensional pan–tilt and piezoelectric micromotion platform [42]; its structure is shown in Fig. 12.3. The coarse tracking structure adopts a two-dimensional PTZ TGMZ-01 with an azimuth adjustment range of 0°–360° and a pitch range of −70°–70°. The angle adjustment in the two directions is controlled by a servomotor, and the control algorithm adopts an incremental PID algorithm.

The precision tracking structure adopts a piezoelectric micromotion platform, and its azimuth and pitch angles can vary from 0 mrad to 0.2 mrad, and from 0 mrad to 0.4 mrad. The test results for the coarse tracking system are shown in Fig. 12.4. An incremental PID algorithm was used for control to suppress the positioning error of the system. The error suppression performance of the system was tested when the target angle increment was 0.03°.

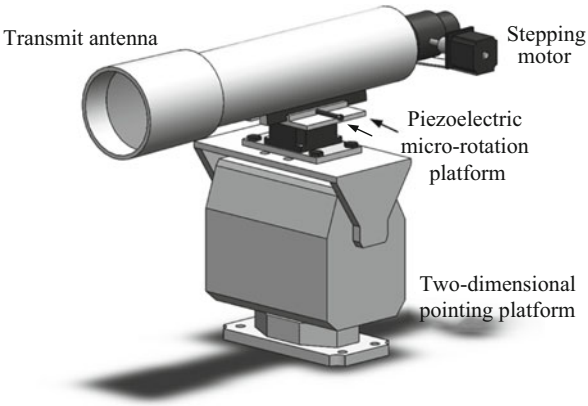


Fig. 12.3 Coarse- and fine-grading control system [42]

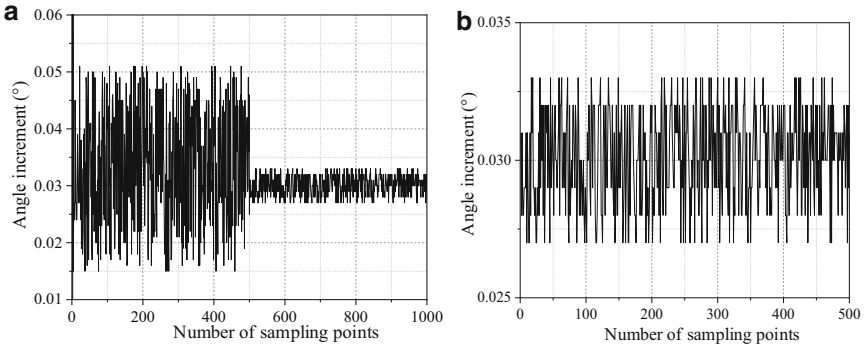


Fig. 12.4 Angle increment (**a**) before and (**b**) after error suppression [43]

Figure 12.5 shows the response curve of the coarse tracking structure. The figure shows that the system has no overshoot or large-amplitude oscillations and that the control performance is ideal.

The relationship curve between the input voltage and angle of the piezoelectric micromotion platform is shown in Fig. 12.6, where Fig. 12.6a is the θ_x relation curve of the voltage and angle in the x direction and Fig. 12.6b is the θ_z curve of the voltage and angle in the z direction. The linearities $\delta = (\Delta y_{\max}/y) \times 100\%$ available in the θ_x and θ_z directions are 0.15% and 0.14%, respectively. The results are approximated

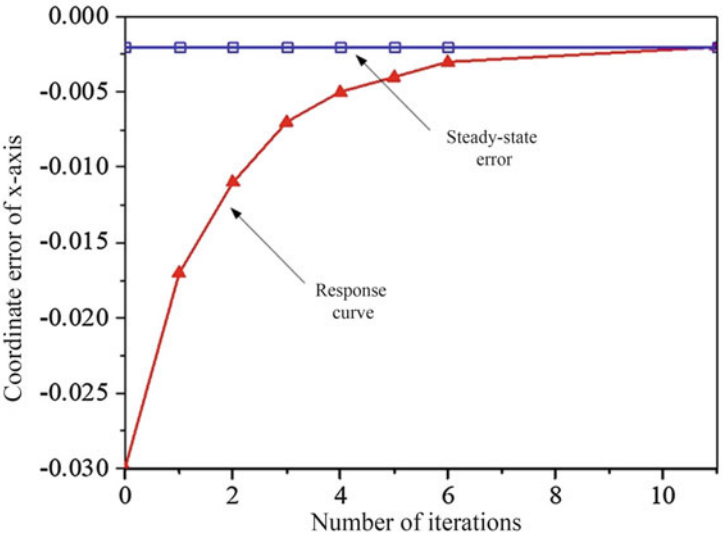


Fig. 12.5 Two-dimensional PTZ response curve [43]

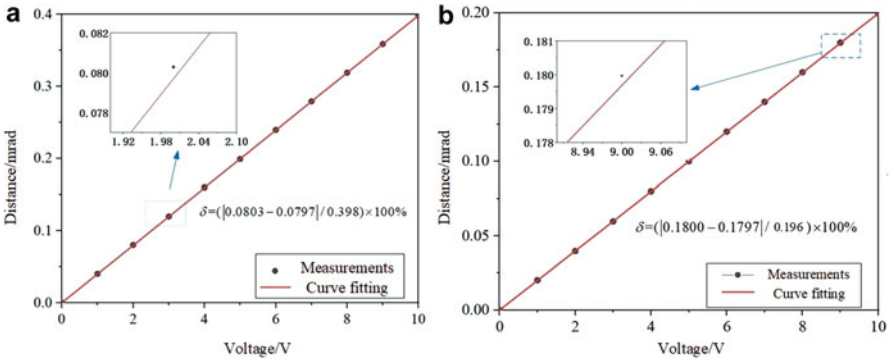


Fig. 12.6 Relation curve between input voltage and angle. (**a**) θ_x-direction control voltage and angle; (**b**) linear error curve of θ_z direction [43]

using a linear relationship. Δy_{max} is the maximum deviation between the input voltage, angular-displacement relationship curve, and ideal fitting curve. y is the full-scale output value [43].

Beam Detection System

The optical path diagram of a beam-aiming detection system is shown in Fig. 12.7. The beam was divided into communication and position detection beams after passing through a light splitting prism. The former was used to transmit information, and the latter was irradiated on the optical rear projection screen to the image.

The four beam detection cases are shown in Fig. 12.8. α represents the angle between the incident beam and the visual axis of the antenna, and ρ represents the beam offset. By defocusing the system, the problem of the system being unable to detect the beam offset can be overcome.

Automatic Tracking Control System for Space-Based Relay Optical Wireless Communication

Figure 12.9 shows the schematic diagram of a UAV-relay APT system with an integrated transceiver. The six-rotor UAV is equipped with a two-dimensional reflector and an optical receiving unit to capture, aim, and track the signals sent from the ground terminal, to enable the two ground terminals to realize two-way communication through the UAV relay platform. When ground terminal 1 is the transmitting terminal, ground terminal 2 is the receiving terminal, and vice versa. Thus, UAV-relay laser communication with an integrated transceiver can be realized [44].

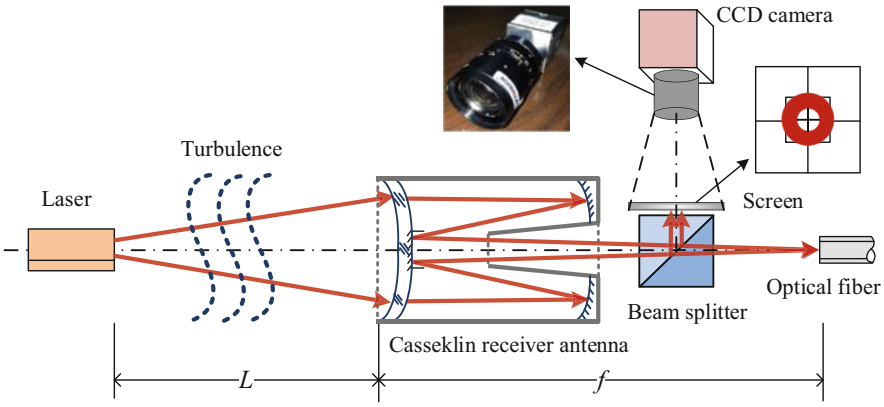


Fig. 12.7 Schematic diagram of the beam detection system structure [5]

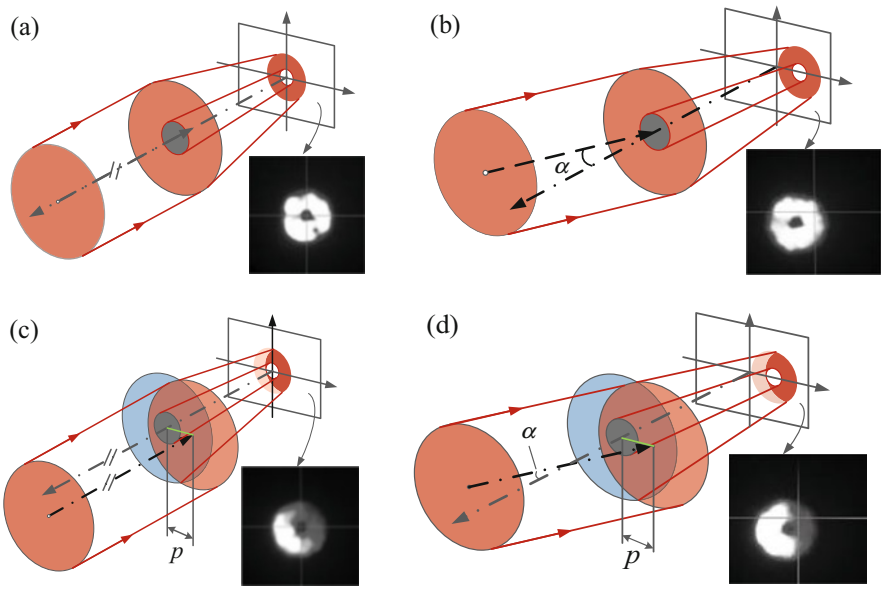


Fig. 12.8 Four beam detection cases [5]. (**a**) $\alpha = 0, p = 0$; (**b**) $\alpha \neq 0, p = 0$; (**c**) $\alpha = 0, p \neq 0$; (**d**) $\alpha \neq 0, p \neq 0$

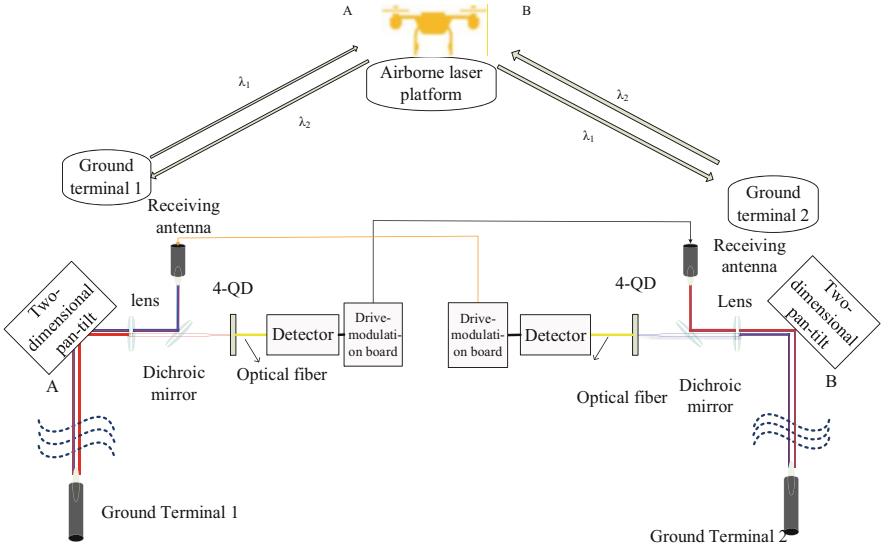


Fig. 12.9 Integrated transceiver UAV-relay APT system [45]

In the experiment, the signal wavelength was 650 nm, transmitted optical power was 2 mW, and communication rate was 10 Mbit/s. The transmitting end used on–off

keying (OOK) modulation to load the video signal onto the light source and collimate the transmission output through a transmitting antenna. The two-dimensional reflector focused the laser beam through the lens onto the through-hole four-quadrant detector to complete the coarse aiming process [46]. The detector then converted the received optical signal into photoelectricity, the signal was demodulated, and the image information collected by the transmitter was observed on the computer.

Figure 12.10 shows the signal waveform at the receiving end. The signal output by the transmitter is a 10-Mbit/s square wave signal. From the waveform of the signal at the receiver, it can be observed that the signal is not significantly distorted and is still a relatively complete square wave signal. The test results show that the system satisfies the performance index for working at 10 Mbit/s.

Figure 12.11 shows the distribution of the spot center coordinate positions during the aiming process. The image was re-aimed after 4500 spot drifts. The re-aimed images of the spot were collected, and the center position coordinates of the spot were obtained through image processing. The black "+" in the figure refers to the position of the bull's eye, and the blue coordinate points refer to the positions of the light spot center deviating from the bull's eye at the time of aiming. After the calculation, the aiming accuracy of the system reached 2.42 μrad.

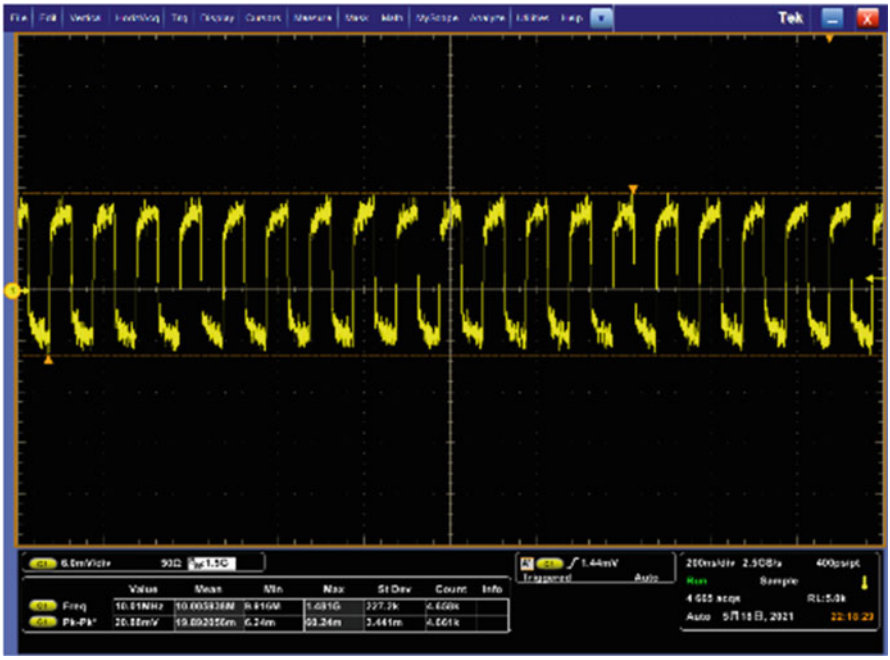


Fig. 12.10 Oscilloscope signal waveform at the receiving end [45]

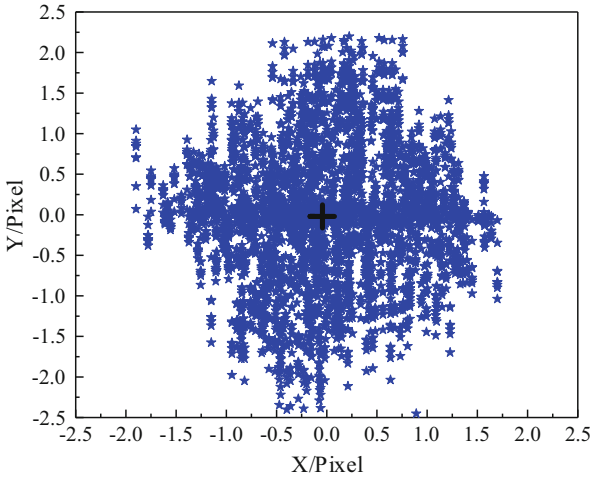


Fig. 12.11 Spot center coordinate position distribution [45]

Field Experiments for Acquisition, Aiming, and Tracking in Optical Wireless Communication

1.3-km Far-Field Experiment

The assembly structure diagram of the 1.3-km far-field experimental system is shown in Fig. 12.12. The transmitting end was located on the eighth floor of the teaching building on the Jinhua Campus of Xi'an University of Technology. The receiving end was located on the 11th floor [5] of Kaisen Fujing Yayuan, East 2nd Ring Road, Xi'an, China, with a communication link distance of 1.3 km. The experimental system was based on a NexStar 8SE astronomical telescope. A camera was placed before and after the receiving antenna for capturing and tracking, and the entire system was placed on a single-arm platform.

The spot data measured in the experiment are shown in Fig. 12.13. In the light spot tracking process, the light spot drifted out of the tracking threshold 11 and two times in the azimuth and pitch directions, respectively. After the system was adjusted, the light spot returned to its coordinate origin [43].

5.2-km Far-Field Experiment

A schematic diagram of the 5.2-km far-field experiment system is shown in Fig. 12.14. The computer processed the spot image information collected from the CCD camera and obtained the spot position coordinates. After receiving the data, the transmitter directed the optical antenna according to the spot position error information to achieve stable beam tracking.

Figure 12.15 shows the spot data measured during the tests. Figure 12.16a and b show the drifts of the spot positions in the azimuth and pitch directions, respectively, with a data sampling time of 12 h. Figures 12.17a and b show the statistical data of the spot position error measured in the test in the azimuth and pitch directions, respectively.

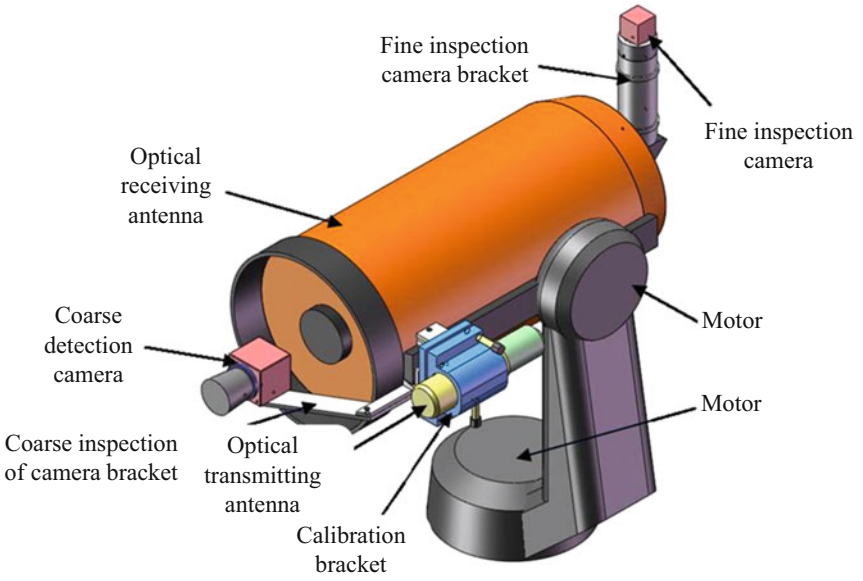


Fig. 12.12 Assembly structure diagram of the 1.3-km far-field experimental system [5]

As shown in Fig. 12.16, from the results of the light spot drift, the light spot in the azimuth and pitch directions drifted out of the threshold 23 and 14 times, respectively. The light spot returned to the center point of the CCD camera after adjustment by the tracking system.

10.2-km Far-Field Experiment

The 10.2-km far-field experimental system assembly is shown in Fig. 12.17; it is composed of a laser light source, optical fiber amplifier, transmitting antenna, two-dimensional aiming platform, and its driver. The receiving end is composed of a receiving antenna, infrared CCD camera, and its controller.

During the experiment, the spot image was sampled every 1 min, and the spot position at the sampling time was obtained after image processing. The azimuth and pitch deviations of the spot refer to the difference between the spot position centers in the two different states.

The detection curves of the light spot images under different weather conditions are shown in Figs. 12.18, 12.19, and 12.20. The results show that the range of spot offsets is 50–110 pixels [48].

Figure 12.21a and b show the fitting curves of the light spot center during the experiment [48]. The results show that the center position coordinates of the light spot change with different external environment factors, such as temperature, humidity, and wind speed.

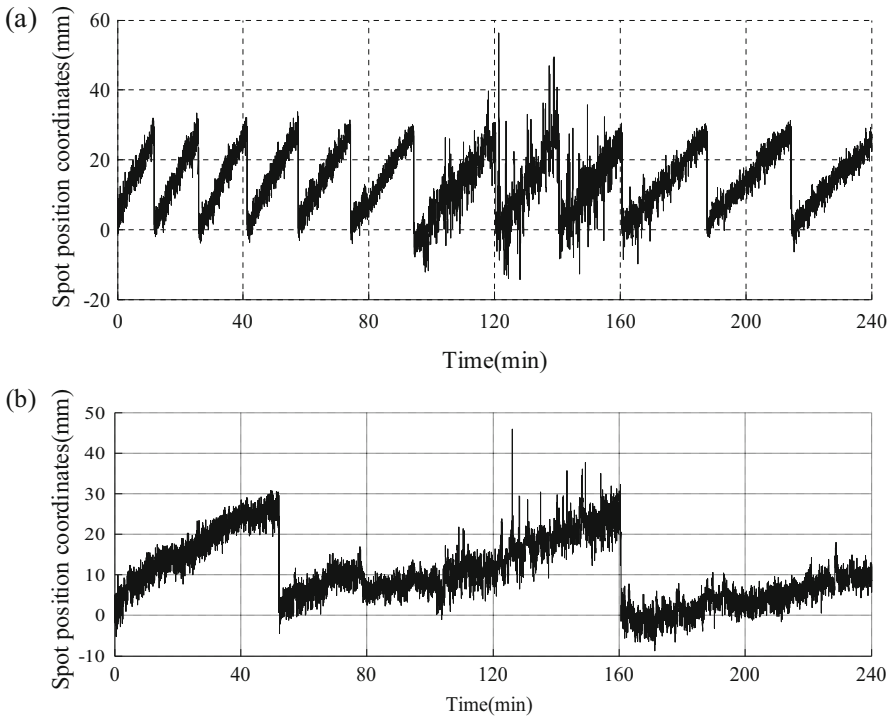


Fig. 12.13 Spot position coordinates (2016.5.25 22:46–2016.5.26 22:00; light rain, 13 °C–18 °C). (**a**) Azimuth direction; (**b**) pitch direction [5]

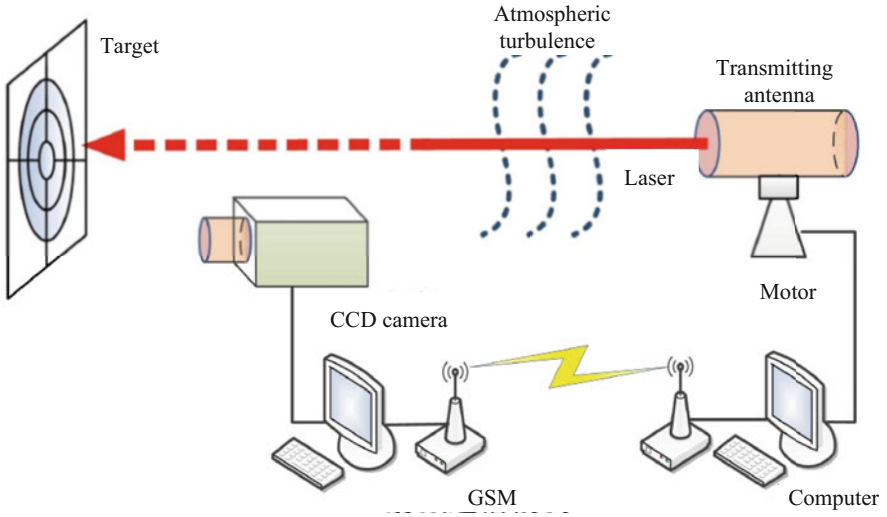


Fig. 12.14 Schematic diagram of the spot tracking system [47]

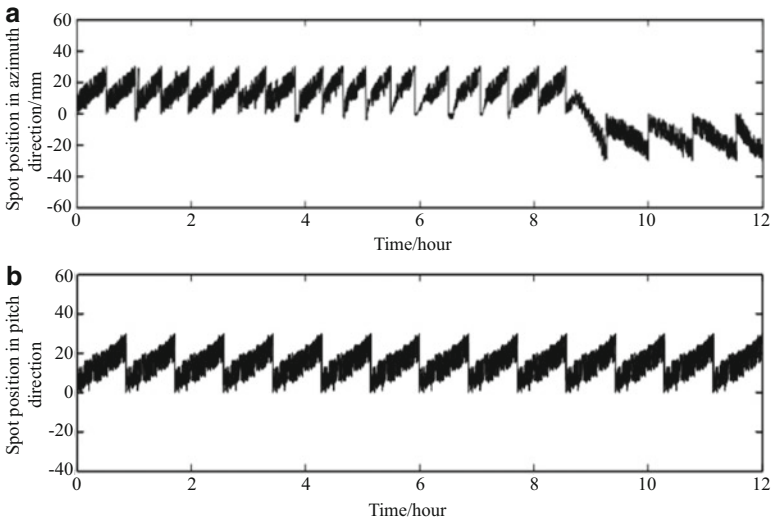


Fig. 12.15 Spot tracking curve (2018.13.5 18:00–2018.13.6 6:00; cloudy, −1°–9°) [47]. (**a**) Azimuth; (**b**) pitch direction

100-km Far-Field Experiment

The system assembly photos of the 100-km far-field experiment are shown in Fig. 12.22. The experiment adopted a coarse- and fine-grading control system. The transmitter was located in Erlangjian Scenic Spot, Gonghe County, on the south bank of Qinghai Lake, at an altitude of 4015 m. The receiving end was located in Quanji Township, Gangcha County, on the north bank of Qinghai Lake, at an altitude of 3955 m [43]. The optical wavelength of the signal was 1550 nm and the system adopted a double-transmitter and single-receiver structure.

The 100-km outfield experiment used the rectangular scanning method to capture the beam, and the aiming and tracking process was similar to that in the 10.2-km far-field experiment. The change curve of the spot center position is shown in Figs. 12.23 and 12.24. It can be seen from the change of the spot center position that the tracking and aiming of the system is more difficult, owing to the influence of atmospheric turbulence and the increase in communication link distance.

Non-common Visual Axis Two-Dimensional Mirror Rapid Aiming System

Basic Principle

Figure 12.25 shows the capture, alignment, and tracking system controlled by the transmitter image calibration and receiver using a two-dimensional mirror. The coarse alignment calibration from the transmitting antenna to the two-dimensional

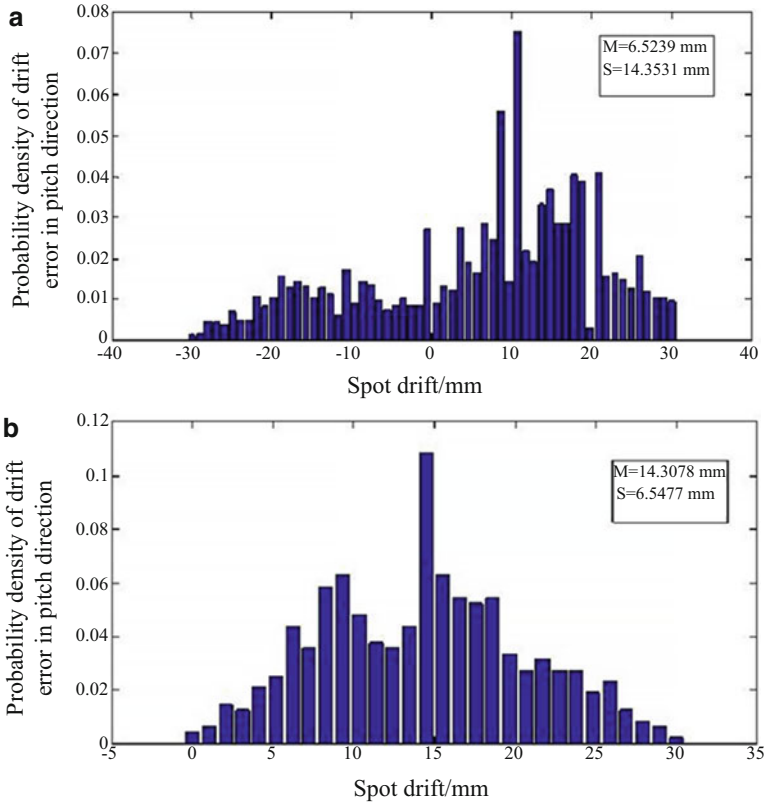


Fig. 12.16 Spot position error statistics (2018.13.5 18:00–2018.13.6 6:00; cloudy, −1°–9°). (**a**) Azimuth; (**b**) pitch direction

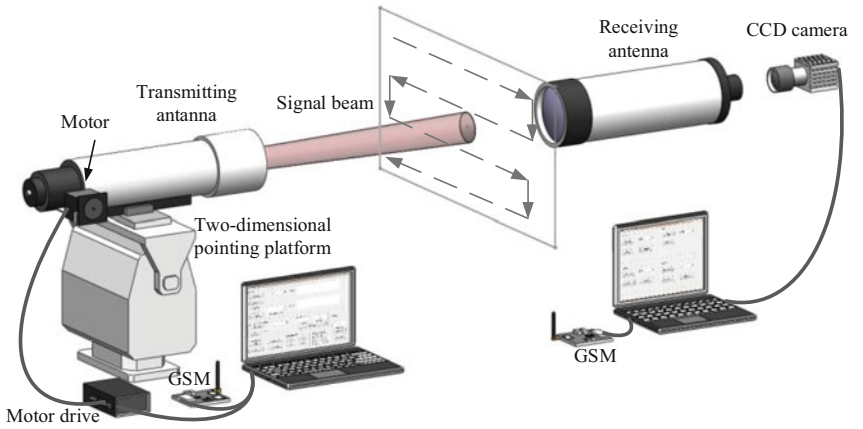


Fig. 12.17 Diagram of the 10.2-km beacon-free optical APT system assembly [43]

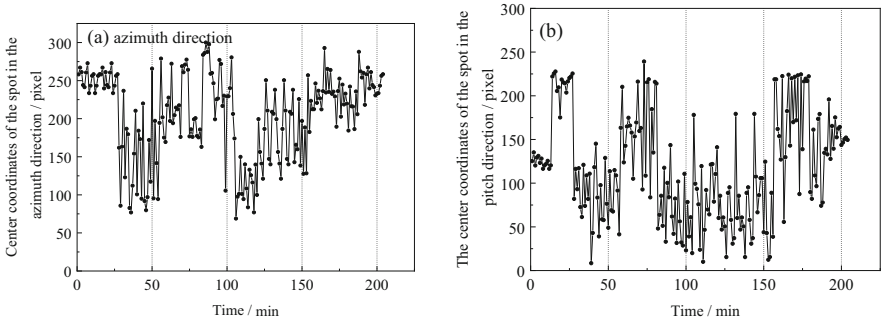


Fig. 12.18 Spot center coordinate curve (2019.10.30 21:00–2019.11.1 00:00; sunny, 17 °C). (**a**) Azimuth direction; (**b**) pitch direction [48]

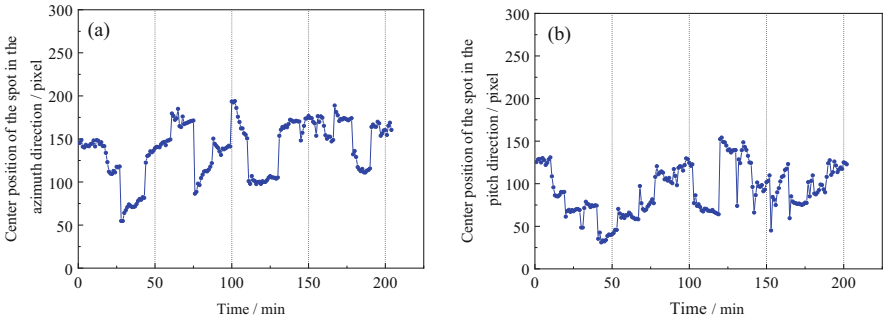


Fig. 12.19 Spot center coordinate curve (2019.11.1 21:00–2019.11.2 0:00; negative, 12 °C). (**a**) Azimuth direction; (**b**) pitch direction [48]

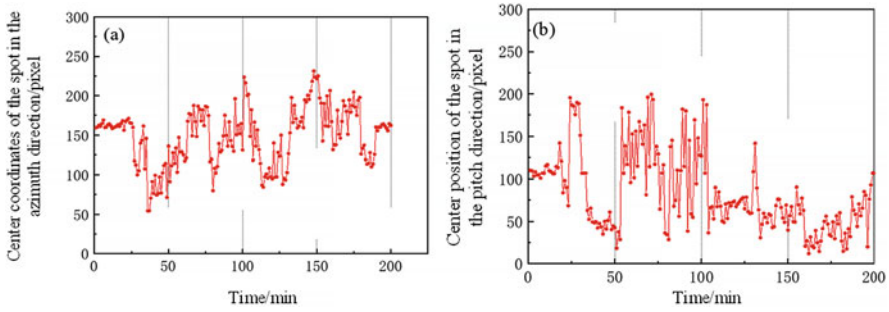


Fig. 12.20 Spot center coordinate curve (2019.11.2 21:00–2019.11.2 0:00; cloudy, 14 °C). (**a**) Azimuth direction; (**b**) pitch direction [48]

reflector uses the calibration camera at the transmitting end to point to a designated area at the receiving end, directly calibrates the position, and adjusts the transmitting antenna with the image as feedback.

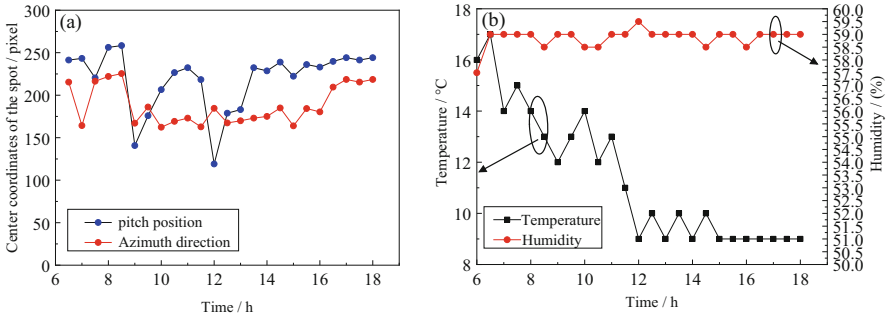


Fig. 12.21 (**a**) Spot center position change curve; (**b**) temperature and humidity change curve [48]



Fig. 12.22 The 100-km outfield experiment scenario. (**a**) Receiver; (**b**) transmitter [48]

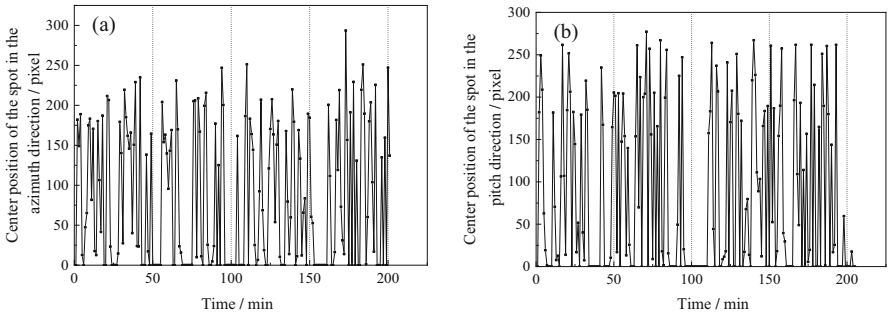


Fig. 12.23 Change curve of spot center coordinates (first experiment) (20110.9.18 23: 00–20110.9.19 02:00; sunny, 14 °C). (**a**) Detection of spot center coordinates in azimuth direction; (**b**) detection of spot center coordinate in pitch direction [58]

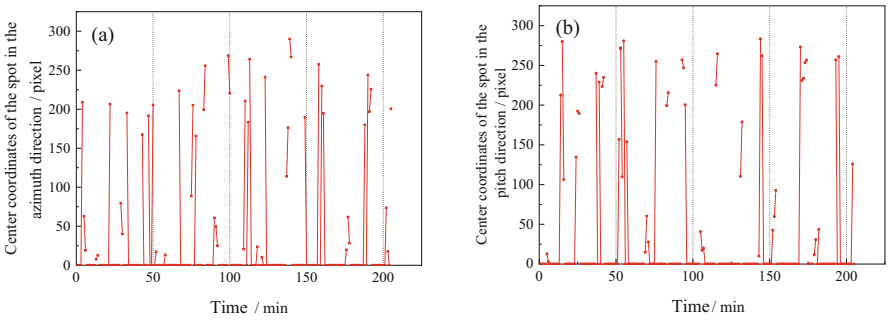


Fig. 12.24 Change curve of spot center coordinates (second experiment) (20110.9.20 23: 00–20110.9.20 02:00; overcast to light rain, 9 °C). (**a**) Detection of spot center coordinates in azimuth direction; (**b**) detection of spot center coordinates in pitch direction [48]

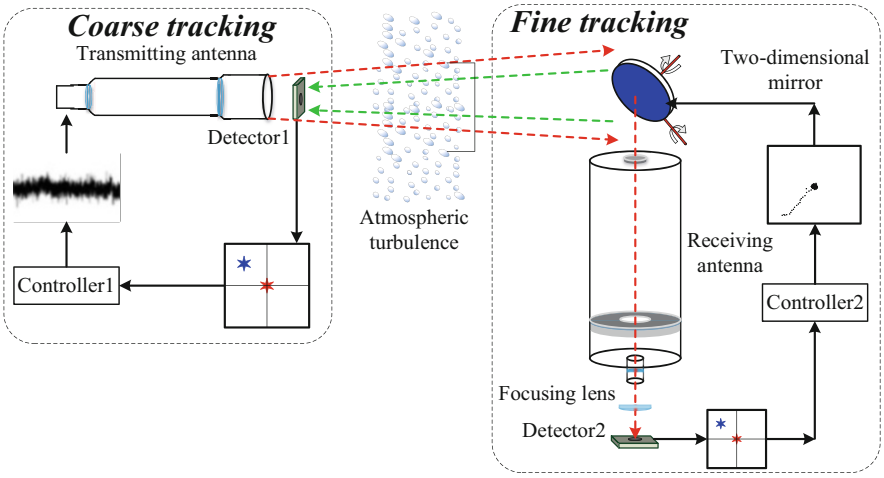


Fig. 12.25 Image calibration at the transmitter; acquisition, alignment, and tracking system controlled by the two-dimensional mirror at the receiver [50]

The short-axis fine alignment from the two-dimensional mirror to the receiving antenna adjusts the two-dimensional mirror according to the feedback information from the detector at the rear of the receiving antenna. The model can be operated at either end, and no auxiliary communication is required for data interaction at the receiving or sending ends. This alignment method is convenient and less affected by the environment.

A beam acquisition, alignment, and tracking system with a non-common axis of view is used to quickly establish the optical wireless communication link, which is divided into three stages: acquisition (image calibration at the transmitter), alignment, and tracking.

1. Acquisition stage: The transmitter gradually approaches the two-dimensional reflector of the receiver by adjusting the azimuth and pitch angles of the transmitter antenna and calibrating the camera imaging. It is only necessary to find the target point to combine the center of the camera's field of view with the center of the image target of the captured long-range two-dimensional reflector, which completes the long-distance beam alignment.
2. Alignment stage: After the beam is aligned with the two-dimensional mirror at the receiving end, the two-dimensional mirror scans the beam within a certain range by controlling the stepper motor. When the optical signal is captured by the spot detection unit at the receiving end, the two-dimensional mirror stops scanning. Then, the position information of the spot is detected and calculated by computer software to adjust the two-dimensional mirror, thus gradually approaching the position of the reference point.
3. Tracking stage: The error is corrected in real time according to the beam position detection results. When a light spot drifts out of the preset threshold range, the position deviation is corrected. If the external interference suddenly increases, such that the system fails to track and the spot target is lost, the two-dimensional mirror scans again until the spot target appears in the field of view of the detector.

Coarse Tracking and Control Model

In Fig. 12.26, the coarse tracking at the transmitter uses a zoom camera as the image detection device, and the calibration principle is shown in Fig. 12.28. The zoom camera at the transmitting end is located at the position of the image in the Fig. 12.26, whereas the calibrated object at the receiving end is located at the position of the object in the picture.

Assuming that when the zoom capacity of the camera reaches its maximum (i.e., when the maximum focal length is reached), if the calibrated object can present an image that can be resolved on a single pixel point, the observation distance is at its limit. Assuming that the size of a pixel in the calibration camera is x_{pixel}, the maximum zoom length of the camera is f, and the size of the calibration object at the receiving end is d, the maximum limit distance L that the camera can calibrate can be expressed as

$$L = \frac{f}{x_{\text{pixel}}} \cdot D. \tag{12.1}$$

When the camera resolution x_{pixel}, zoom ability f, and other limitations make it impossible for the calibration camera to observe calibration point 1 in Fig. 12.27, the

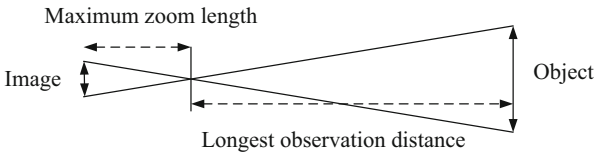


Fig. 12.26 Coarse tracking mark imaging model [50]

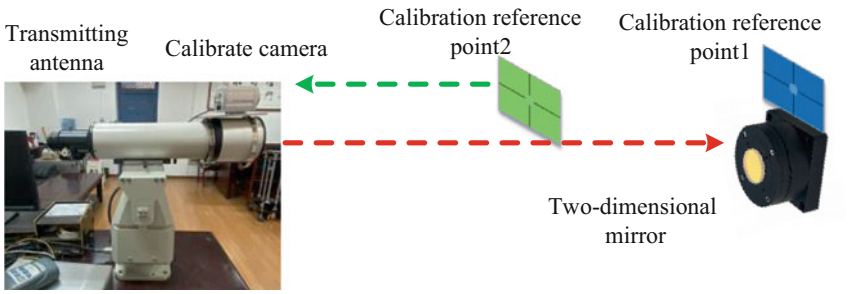


Fig. 12.27 Schematic diagram of the calibration principle using a coaxial reference system [50]

traditional alignment method can be used to align the beam. When the beam is aligned, calibration point 2 should be placed at the limit observation distance close to the camera and used as the reference point for the coarse alignment and tracking of the beam.

Because the calibration camera and optical antenna are mechanical structures with optical axes parallel to each other, calibration point 2 can also be aligned and tracked after calibration point 1 to coarse align and track the beam to the two-dimensional mirror. When the observation distance limit of the calibration camera is insufficient, a near-point coaxial reference system can be selected.

Two-Dimensional Mirror Model

After coarse aligning the beam, the spatial coordinate system shown in Fig. 12.28 is established to study the scanning track of the beam on the exit surface after passing through the two-dimensional reflector.

Assume that the vector of the incident ray is recorded as A, and the angle between the incident ray and the normal of the two-dimensional reflector is α. The vector direction of the incident ray is considered as the x' axis, and the spatial coordinate system of the incident ray is established (x', y', and z'). The vector of outgoing light is marked as A'.

Considering the vector direction of the outgoing ray as the z axis, the spatial coordinate system (x, y, and z) of the outgoing ray is established. The normal vector of the mirror is N, and the deflection angles of the two-dimensional mirror in the transverse and longitudinal directions are θ_1 and θ_2, respectively.

With the reflection ray coordinate system (x, y, and z) as the main coordinate system, vectors A and N are expressed as:

$$\mathbf{A} = \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = \begin{bmatrix} -\sin(2\alpha) \\ 0 \\ -\cos(2\alpha) \end{bmatrix} \tag{12.2}$$

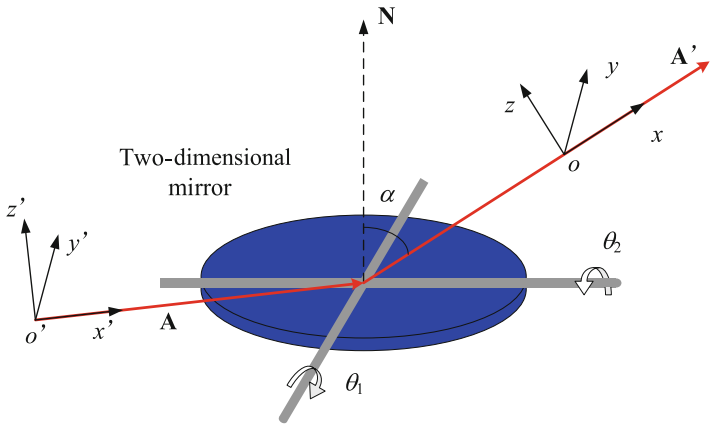


Fig. 12.28 Geometric optical model of the two-dimensional mirror [50]

$$\mathbf{N} = \begin{bmatrix} N_x \\ N_y \\ N_z \end{bmatrix} = \begin{bmatrix} \sin(\alpha + \theta_1) \cdot \cos\theta_2 \\ \sin\theta_2 \\ \cos(\alpha + \theta_1) \cdot \cos\theta_2 \end{bmatrix}. \tag{12.3}$$

According to the reflection law, each vector is expressed as:

$$\mathbf{A}' = \mathbf{A} - 2(\mathbf{A} \cdot \mathbf{N}) \cdot \mathbf{N} = \mathbf{MA} \tag{12.4}$$

$$\mathbf{N}(\theta) = \mathbf{N} \cdot \cos\theta + \mathbf{C}(\mathbf{C} \cdot \mathbf{N}) \cdot (1 - \cos\theta) + (\mathbf{C} \times \mathbf{N}) \cdot \sin\theta. \tag{12.5}$$

In Eq. (12.5), C denotes the unit direction vector corresponding to any rotational axis direction. Equations (12.2), (12.3), and (12.4) are substituted into Eq. (12.5) to obtain the matrix representation of M:

$$\mathbf{M} = \begin{bmatrix} 1 - 2N_x^2 & -2N_xN_y & -2N_xN_z \\ -2N_xN_y & 1 - 2N_y^2 & -2N_yN_z \\ -2N_xN_z & -2N_yN_z & 1 - 2N_z^2 \end{bmatrix}. \tag{12.6}$$

The outgoing-ray matrix is expressed as:

$$\mathbf{A}' = \begin{bmatrix} A'_x \\ A'_y \\ A'_z \end{bmatrix} = \begin{bmatrix} \sin(2\alpha) \cdot \left(2 \cdot \cos^2\theta_2 \cdot \sin^2(\alpha + \theta_1) - 1\right) + \cos(2\alpha) \cdot \cos^2\theta_2 \cdot \sin(2\alpha + 2\theta_1) \\ \cos(2\alpha) \cdot \cos(\alpha + \theta_1) \cdot \sin(2\theta_2) + \sin(2\alpha) \cdot \sin(\alpha + \theta_1) \cdot \sin(2\theta_2) \\ \cos(2\alpha) \cdot \left(2 \cdot \cos^2\theta_2 \cdot \cos^2(\alpha + \theta_1) - 1\right) + \sin(2\alpha) \cdot \cos^2\theta_2 \cdot \sin(2\alpha + 2\theta_1) \end{bmatrix}. \tag{12.7}$$

Therefore, the coordinates of the two points on the scanning path of the light spot are expressed as:

$$x = L\frac{A'_x}{A'_z} = L\frac{\sin(2\alpha)\cdot\left(2\cdot\cos^2\theta_2\cdot\sin^2(\alpha+\theta_1)-1\right)+\cos(2\alpha)\cdot\cos^2\theta_2\cdot\sin(2\alpha+2\theta_1)}{\cos(2\alpha)\cdot(2\cdot\cos^2\theta_2\cdot\cos^2(\alpha+\theta_1)-1)+\sin(2\alpha)\cdot\cos^2\theta_2\cdot\sin(2\alpha+2\theta_1)} \tag{12.8}$$

$$y = L\frac{A'_x}{A'_z} = L\frac{\cos(2\alpha)\cdot\cos(\alpha+\theta_1)\cdot\sin(2\theta_2)+\sin(2\alpha)\cdot\sin(\alpha+\theta_1)\cdot\sin(2\theta_2)}{\cos(2\alpha)\cdot(2\cdot\cos^2\theta_2\cdot\cos^2(\alpha+\theta_1)-1)+\sin(2\alpha)\cdot\cos^2\theta_2\cdot\sin(2\alpha+2\theta_1)}. \tag{12.9}$$

In Eqs. (12.8) and (12.9), L is the linear distance from the beam reflection point to the *xoy* plane.

According to Eqs. (12.8) and (12.9), let $L = 1$ m describe the scanning track of the two-dimensional mirror under different incident angles α, and different scanning angles θ_1 and θ_2, as shown in Fig. 12.29. From Fig. 12.29a and b, it can be seen that, under the same scanning parameters in the two directions of scanning angles θ_1 and θ_2, the greater the incident angle α, the greater the radian of the track curve scanned by the two-dimensional mirror.

Compared with Fig. 12.29c and d, when the incident angle α is the same, the smaller the scanning angles θ_1 and θ_2, the larger the area of the beam scanning path. The value of the incident angle α determines the radian of the scanning path of the two-dimensional mirror to the beam, whereas the parameters of the scanning angles θ_1 and θ_2 determine the size of the scanning area.

Therefore, to make the transmitting antenna and receiving antenna realize non-collinear axis alignment in the two-dimensional reflector spot tracking system, we should select reasonable scanning angles θ_1 and θ_2 for the two-dimensional reflector, according to the incident angle of the beam on the two-dimensional reflector surface of the receiving end in the actual field experiment.

Fine Tracking and Its Control Model

As shown in Fig. 12.30, fine beam tracking can be achieved using the classical integral control. Taking the deviation of the actual spot from the position of the reference center point as feedback, the displacement value is converted into the angle-driving value of the motor to track the beam. The k^{th} iteration can be expressed as:

$$\begin{bmatrix}\theta_1(k+1)\\ \theta_2(k+1)\end{bmatrix} = \begin{bmatrix}m(k)-m_c\\ n(k)-n_c\end{bmatrix}\cdot T\cdot k_{\text{i}} + \begin{bmatrix}\theta_1(k)\\ \theta_2(k)\end{bmatrix}. \tag{12.10}$$

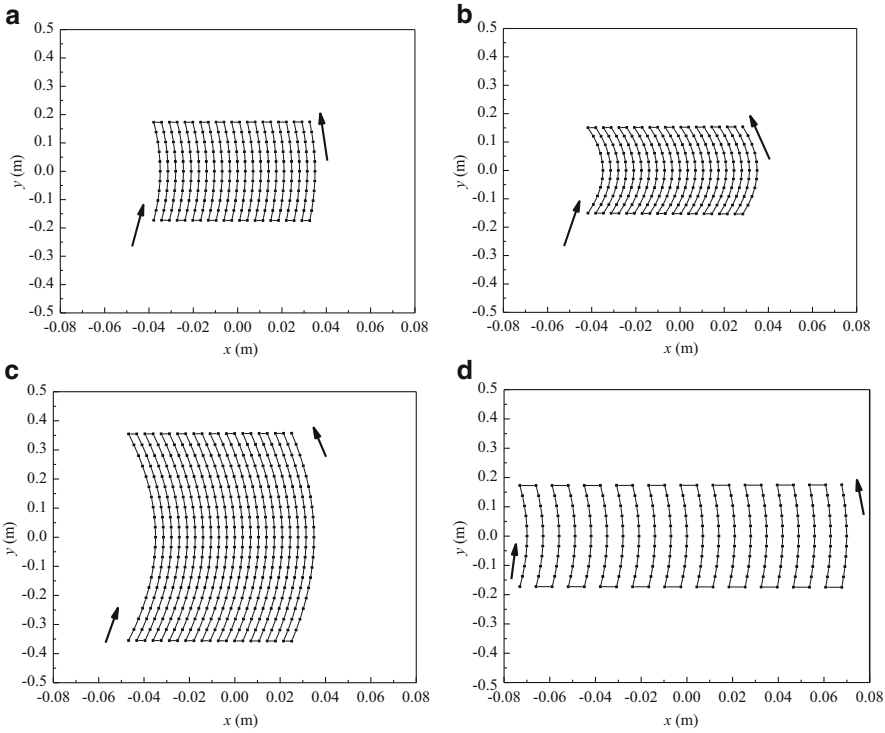


Fig. 12.29 Exit light spot trajectory under different incident angles and scanning angles [50]. (**a**) $\alpha = 15°$, $\theta_1 = -2°–2°$, $\theta_2 = -6°–6°$; (**b**) $\alpha = 30°$, $\theta_1 = -2°–2°$, $\theta_2 = -6°–6°$; (**c**) $\alpha = 15°$, $\theta_1 = -2°–2°$, $\theta_2 = -12°–12°$; (**d**) $\alpha = 15°$, $\theta_1 = -4°–4°$, $\theta_2 = -6°–6°$

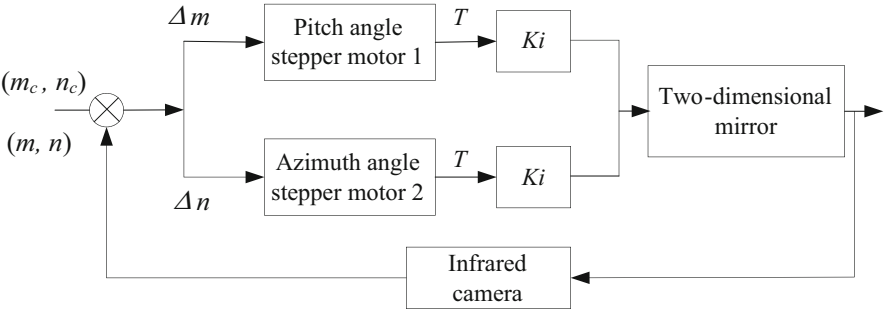


Fig. 12.30 Schematic diagram of the beam tracking algorithm using a two-dimensional mirror [50]

In Eq. (12.10), T is the proportion coefficient from the unit pixel displacement of the infrared camera to the unit step angle of the stepping motor, and k_i is the integral coefficient. (m_c, n_c) is the position of the central reference point of the detector,

$[m(k), n(k)]$ is the actual spot-imaging position, and $[\theta_1(k), \theta_2(k)]$ is the offset angle of the pitch and azimuth angles of the stepper motor, relative to the initial position in the k^{th} iteration.

A far-field experimental communication link using this system is shown in Fig. 12.31. The transmitting and receiving ends were located on the sixth teaching floor in Xiaozhai Village, Bailuyuan, Xi'an City, and Jinhua Campus, Xi'an University of Technology, respectively, with a communication link length of 11.3 km.

As shown in Fig. 12.32, the beam tracking curve of the position coordinates of the two-dimensional mirror is shown, where X represents the pitch direction, Y represents the azimuth direction, and the two-dimensional mirror is adjusted four and two times in the pitch and azimuth directions, respectively.

When the light spot gradually moves to the camera center point, the mean value, variance, and power spectral density (PSD) of the change in the light spot center



Fig. 12.31 Diagram of the 10.2-km optical wireless communication experimental link [50]

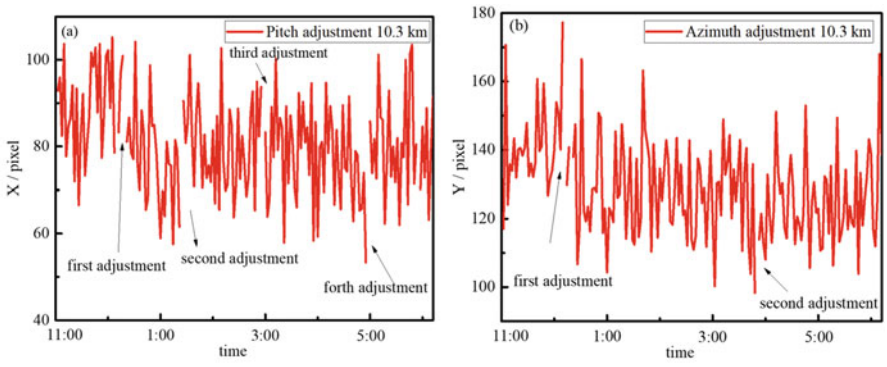


Fig. 12.32 Beam tracking curve (2021.8.24 23:00–2021.8.25 6:00). (**a**) Pitch direction; (**b**) azimuth direction [50]

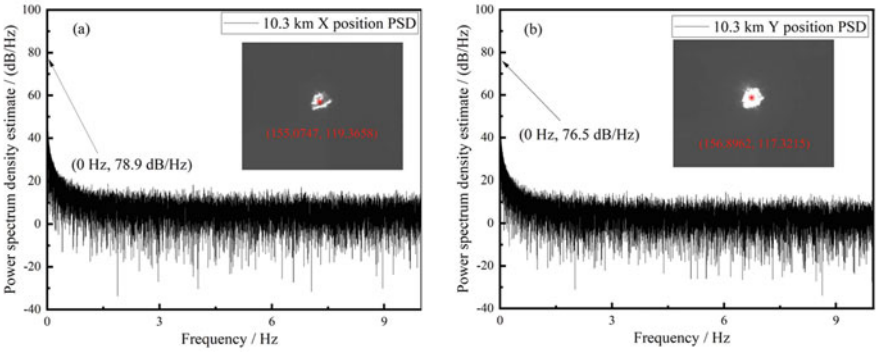


Fig. 12.33 Power-spectral density estimation. (**a**) *X* direction; (**b**) *Y* direction [50]

coordinate in the pitch and azimuth directions are calculated. The results are presented in Fig. 12.33. From the results, it can be observed that the drift in the center position coordinate of the light spot in the azimuth direction is greater than that in the pitch direction.

Figure 12.34 shows the motion track of the light spot and the corresponding PSD during the tracking process. It can be seen from the results that the variance of the light spot in the pitch and azimuth directions is 16.4970 pixel^2 and 9.0287 pixel^2, respectively.

Figure 12.35 shows the signal waveforms at both ends of the transceiver. The transmitter power is 200 mW. When the tracking process is completed, the signal waveform output by the detector for the photoelectric conversion is 74.4 mV.

Chapter Review

This chapter summarized the research progress of Xi'an University of Technology in APT and introduced the working principles and experimental analyses of the initial acquisition system, coarse- and fine-grading tracking system, beam detection system, space-based relay optical wireless communication automatic tracking control system, and receiving and transmitting non-coaxial two-dimensional reflector rapid aiming system. The application scenarios are summarized as follows.

Sections 12.3.1–12.3.3 respectively described the initial acquisition system, coarse- and fine-grading control system, and beam detection system, according to the actual requirements of each stage of acquisition, aiming, and tracking. The initial acquisition system rapidly acquired the position through position calibration, without the need for the positioning devices in traditional acquisition systems, thereby reducing the acquisition time and system design costs.

The coarse- and fine-grading control system used a two-dimensional pan–tilt + piezoelectric micromotion platform as the tracking and aiming actuator. It suppressed the aiming and tracking residual errors not eliminated by the coarse

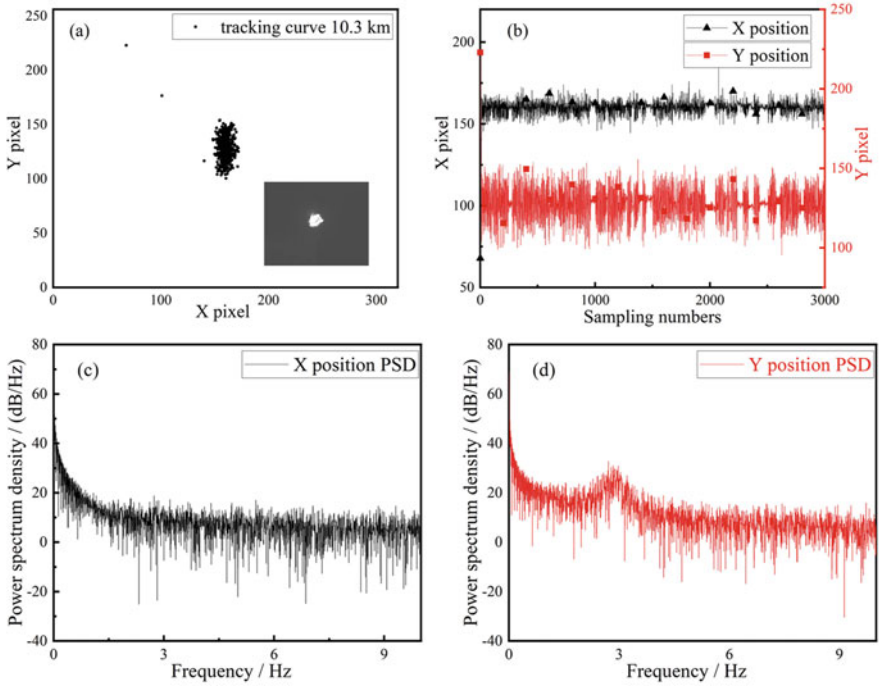


Fig. 12.34 Spot tracking curve and PSD. (**a**) Spot tracking curve; (**b**) direction-tracking curve; (**c**) power spectral density estimation in the *X* direction; (**d**) power spectral density estimation in *Y* direction [50]

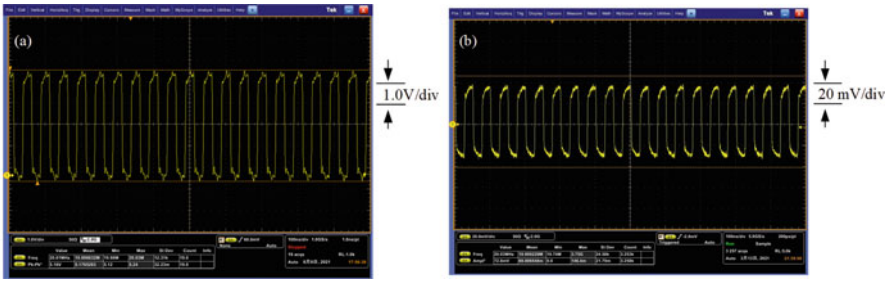


Fig. 12.35 Signal waveform at both ends of the transceiver. (**a**) Transmitting signal; (**b**) receiving signal [50]

tracking and aiming mechanism, and further improved the tracking and aiming accuracy. Its application scenarios were described in detail in section 12.3.5.4. The beam detection system could calculate the beam deviation in the pitch and azimuth directions during the test using data such as the centroid and spot centroid.

The transmitter of the rapid aiming system of the non-coaxial two-dimensional reflector was calibrated by the camera. The receiver scanned the beam by

controlling the two-dimensional reflector and tracked the beam with the spot position information as the feedback signal. Traditional long-axis collimation transmits the positional information of a light spot at a detector over a long distance, which is easily affected by atmospheric turbulence. This system used the image as feedback information to adjust the angle of the transmitting antenna without data transmission. It achieved rapid aiming with little influence from the external environment.

From the optical wireless far-field communication experiments at different distances, it could be observed that, with an increase in the communication link distance, the requirements for the servo bandwidth, degrees of freedom, and resolution in the pitch and azimuth directions of the APT system were higher, which is also the main reason for limiting the working distance of the APT system.

For example, in the 10.2-km far-field experiment, the transmitter completed the aiming and tracking using only the two-dimensional aiming pan–tilt as the actuator, and its resolution was approximately 18.45 μrad. In the 100-km far-field experiment, to meet the corresponding accuracy requirements, the coarse- and fine-grading control system composed of a two-dimensional pan–tilt and piezoelectric micromotion platform was used to complete the aiming and tracking. The resolution of the piezoelectric micromotion platform was about 0.01 μrad.

Summary

With the progress in science and technology, increasing demand for high-speed and large-capacity communication, and improvement in the level of optics, information science, control science, and other disciplines closely related to optical wireless communication, the development direction of optical wireless communication APT systems mainly includes the following aspects:

1. APT terminals with high integration, low power consumption, small size, and lightweight.
2. The control system should adopt a high closed-loop bandwidth to overcome the impact of mechanical vibration generated by the platform equipped with the APT system on beam tracking.
3. Coarse and fine composite axis technology should be adopted to achieve a high dynamic range and high-precision acquisition and tracking.
4. Based on research on point-to-point optical wireless communication structures, an optical wireless communication network between the transmitter and multiple receiving targets should be constructed, and an optical wireless communication APT system that meets the point-to-point requirements should be designed.
5. In a space laser communication link, to obtain the position of the receiving end quickly and accurately without prior information, the transmitting and receiving ends can use scanning methods to search for and capture the uncertain area. When the spot detection camera at the receiving end detects the spot position, the acquisition is considered successful. Using the corresponding control algorithm,

the angle and angular velocity changes of the two-dimensional PTZ are controlled, and the centroid of the spot is maintained at the center of the camera's field of view to achieve fast and accurate target tracking.

This chapter systematically analyzed the development and application of APT systems in optical wireless communication and introduced the research progress and achievements of Xi'an University of Technology in this field. It included an experimental analysis and verification of the performance of the designed initial acquisition system, coarse- and fine-grading control system, and beam detection system. The experimental results showed that the system could effectively and rapidly acquire and aim at the target. The tracking and aiming accuracies of the system were greatly improved, and the effectiveness and reliability of the APT system were also improved.

References

1. Xizheng Ke, Ke Dong, Optical Wireless Communication [M]. Beijing: Springer Press, 2022.
2. Yan L S, Wang F, Wu W, et al. Current status and key technologies of unmanned aerial vehicle laser communications payloads [J]. Laser and Optoelectronics Progress, 2016, 53(8): 40–49.
3. Mao Y C, Wang H Q, Zhang Y, et al. Research status and development of optical spatial modulation technology [J]. Opto-Electron Eng, 2020, 47(3): 190713.
4. Liu Z J, Shu B H, Wang Y Z, et al. Design for automatically aligning and stabilizing high energy laser beam [J]. Optical Technique, 1999, 01(01): 19–20.
5. Yang P S. Research and design of acquisition pointing tracking system for free-space optical communication [D]. Xi'an: Xi'an University of Technology, 2016.
6. Nakagawa K, Yamamoto A, Toyoda M. Performance test result of LUCE (laser utilizing communications equipment) engineering model [C]// Free-Space Laser Communication Technologies XII, San Jose, California, January 24, 2000: 68–76.
7. Chen C C, Lesh J R. Overview of the optical communications demonstrator [C]// Free-Space Laser Communication Technologies VI, Los Angeles, California, January 23, 1994: 85–94.
8. Arimoto Y, Toyoshima M, Toyoda M, et al. Preliminary result on laser communication experiment using Engineering Test Satellite-VI (ETS-VI) [C]// Free-Space Laser Communication Technologies VII, San Jose, California, February 7–8, 1995: 151–159.
9. Biswas A, Williams G, Wilson K E, et al. Results of the STRV-2 lasercom terminal evaluation tests [C]// Free-Space Laser Communication Technologies X, San Jose, California, January 27–28, 1998: 2–14.
10. Tolker-Nielsen T, Oppenhauser G. In-orbit test result of an operational optical intersatellite link between ARTEMIS and SPOT4 SILEX [C]// Free-Space Laser Communication Technologies XIV, San Jose, California, January 21–22, 2002:1–18.
11. M. Guelman, A. Kogan. et al. Acquisition and Pointing Control for Inter-satellite Laser Communications [J]. IEEE Trans On Aerospace and Electronic Systems, 2004, 40(4): 1239–1248.
12. Boroson D M, Biswas A, Edwards B L. MLCD: Overview of NASA's Mars laser communications demonstration system [C]// Free-Space Laser Communication Technologies XVI, San Jose, California, January 25, 2004: 16–29.
13. Walther F G, Michael S, Parenti R R, et al. Air-to-ground lasercom system demonstration design overview and results summary [C]// Free-Space Laser Communications X, San Diego, California, August 2–3, 2010: 256–264.

14. Schmidt C, Horwath J. Wide-field-of-regard pointing, acquisition and tracking-system for small laser communication terminals [C]// Proceedings of ICSOS, Ajaccio, Corsica, October 12, 2012: 1–18.
15. Moll F, Mitzkus W, Horwath J, et al. Demonstration of high-rate laser communications from fast airborne platform: flight campaign and results [C]. Unmanned/Unattended Sensors and Sensor Networks X, Toulouse, France, September 23–24, 2014: 181–186.
16. Moll F, Horwath J, Shrestha A, et al. Demonstration of High-Rate laser communications from a fast airborne platform [J]. IEEE Journal on Selected Areas in Communications, 2015, 33(9): 1985–1995.
17. Quintana, G. Erry, A. Gomez, et al. Design of a holographic tracking module for long-range retroreflector free-space systems [J]. Applied Optics, 2016, 55(25): 7173–7179.
18. Antonello R, Branz F, Sansone F, et al. High-precision dual-stage pointing mechanism for miniature satellite laser communication terminals [J]. IEEE Transactions on Industrial Electronics, 2020, 68(1): 776–785.
19. Jiang L J, Zhu D W. Designing for 4km terrestrial line sight optical free space communications systems [J]. Optical Communication Technology, 2000, 24(23): 219–222.
20. Luo T, Hu Y. Design and Realization of Laser Beam Acquisition, Pointing and Tracking of Ground Demonstration System for Free Space Optical Communication [J]. Journal of Applied Optics, 2002(02): 14–18.
21. Zhou Y L, Ai Y, Zuo T, et al. Experimentation of realtime acquisition and tracking of dree apace laser beam and analysis of the result [J]. Acta Photonica Sinica, 2005, 34(6): 943–948.
22. Tong S F, Jiang H L, Liu Y Q, et al. Optimum design of bandwidth for the APT coarse tracking assembly in free space laser communication [J]. Opto-Electronic Engineering, 2007, 34(9): 16–20.
23. Zhang L H, Yang Y, Hu S J, et al. The Moving Characteristics of Directional Driving Motor in Space Tracking System [J]. Acta Optica Sinica, 2008(02): 233–238.
24. Pan G F, Zhang J X, Chen J. Common path auto-alignment system [J]. Chinese Journal of Lasers, 2008, 35(10): 1500–1504.
25. Zhang H, Wang R, Jia X, et al. HY-2 satellite and its application [J]. Space International, 2012, 9(01): 5–9.
26. Song Y S, Chang S, Tong S F, et al. Feature analysis of aeronautical laser communication system and airborne laser communication experiment [J]. Chinese Journal of Lasers, 2016, 43(12): 215–226.
27. Dong R, Ai Y. Shan X, et al. Design and communication experiment of fine tracking system for free space optic [J]. Infrared and Laser Engineering, 2012, 41(10): 2718–2722.
28. Qian F, Jia J J, Zhang L, et al.Positioning accuracy of spot-detecting camera in acquisition, tracking, pointing system [J]. Chinese Journal of Lasers,2016, 40(2): 176–182.
29. Meng L X, Zhao D X, Zhang L Z, et al. The test of tracking accuracy and design of airborne laser communication stabilized pod [J]. Acta Armamentarii, 2015, 36(10): 1916–1923.
30. Zhang Y S, Qiu Z A, Guo S, et al. Key Technology Analysis of Airborne Laser Communication System and Its Verification [J]. Electronics Optics and Control, 2017, 24(10): 80–84.
31. Xu J H, Wang X R, Huang Z Q, et al. PID tracking method of space laser communication based on liquid crystal optical phased array [J]. Laser & Optoelectronics Progress, 2017, 54(02): 209–218.
32. Yu Z L, Jiang X M, Cao K R, et al. Hysteresis characteristics of steering mirror driven by piezoelectric actuator and its experimental research [J]. Acta Optica Sinica, 2018,38(08): 383–3810.
33. Cai M H, Kong D C, Tong X G. Research and implementation of coarse fine pointing alignment for single composite shaft detection system [J]. Electro-Optic Technology Application, 2019, 34(1): 63–66.
34. An J X, He X L, Yang Q Y, et al. Research on the application of the air to ground free space optical communication by small UAV [J]. Optical Communication Technology, 2017, 41(06): 10–14.

35. Lu Q, Ren B, Bian J Y. Research on acquisition and tracking technology for the four-quadrant detector [J]. Opto-Electron Eng, 2020, 47(3): 1–6.
36. Li Q. Research on spot position detection technology of space laser communication based on array detector [D]. Beijing: University of Chinese Academy of Science, 2020.
37. Zhong G. Research on beaconless capture and tracking technology of atmospheric optical communication [D]. Chengdu: University of Electronic Science and Technology of China, 2022.
38. Ke X Z, Lu N, Zhan L. A light beam automatic capture device and light beam capture method [P]. China: CN201010185116.10. 2011.
39. Ke X Z, Xi X L, Liu C C. A new auto beam acquiring method for laser communication in atmosphere [J]. Optical Communication Technology, 2004, 28(10): 39–41.
40. Hu Q D. Research on spot detection technology in process of beacon acquisition of laser communication [D]. Xi'an, Xi'an University of Technology, 20010.
41. Ke X Z, Lei S C, Yang P S. Beam coaxial alignment detection in atmospheric laser communication [J]. Chinese Journal of Lasers, 2016, 43(06): 181–190.
42. Zhao Q. Design and implementation of initial acquisition system for wireless laser communication [D]. Xi'an: Xi'an University of Technology, 2016.
43. Xu W. Research and implement of the laser spot detection in ATP system [D]. Xi'an: Xi'an University of Technology, 2016.
44. Ke X Z, Li S Y. Experimental study on optical axis alignment of the optical antenna under the spot broken [J]. Acta Photonica Sinica, 2017, 46(04): 133–141.
45. Ke X Z, Yan X, Yang Yang Y Q, et al. 5.2km Distance wireless laser communication tracking experiment [J]. Computer measurement and control, 2018, 26(11): 233–238.
46. Ke X Z, Jing Y K. Far-field laser spot image detection for use under atmospheric turbulence [J]. Optical Engineering, 2020, 59(1): 1–14.
47. Ke X Z, Zhang P. Automatic Focusing Control in Beaconless APT System [J]. Journal of Russian Laser Research, 2020, 41(1): 61–71.
48. Ke X Z, Liang H L. Airborne laser communication system with automated tracking [J]. International Journal of Optics, 2021, 01(01): 1–9.
49. Yang S J, Ke X Z, Wu J L, et al. Fast alignment of Optical Wireless Communication using two-dimensional mirror [J]. Chinese Journal of Lasers, 2022, 49(11): 101–116.
50. Zhao X. The APT initial capture and error analysis of space laser communication [D]. Changchun: Changchun University of Technology, 2013.
51. Ke X Z, Zhang P. A Optical Wireless Communication tracking control system and tracking control method [P]. China: CN201910339488.10. 2021.
52. Zhang P. Design of acquisition pointing and focusing system for wireless laser communication [D]. Xi'an: Xi'an University of Technology, 2020.
53. Ke X Z, Chen X Z, Wu J L. A UAV relay laser communication system [P]. China: CN201810919713.1. 2019.
54. Liang H L. Design and realization of airborne laser communication automatic tracking control system [D]. Xi'an: Xi'an University of Technology, 2021.
55. Ke X Z, Wang J. A light spot alignment method based on four-quadrant detector [P]. China: CN201611244466.1. 2018.
56. Yan X. Research on spot tracking system in APT system of wireless laser communication [D]. Xi'an: Xi'an University of Technology, Xi'an University of Technology, 20110.
57. Jing Y K. Experimental study on far-field spot image detection for Optical Wireless Communication [D]. Xi'an: Xi'an University of Technology, 2020.
58. Liang J Y, Chen R D, Yao H F, et al. Research progress of acquisition, pointing and tracking in optical wireless communication system [J]. Opto-Electron Eng, 2022, 49(8): 25–48.

Spatial Optical-Fiber Coupling Technology 13

Contents

Introduction 394
Foreign Research Progress 394
Domestic Research Progress 399
Research Progress on Spatial Optical-Fiber Coupling Technology 404
Space-Light–Single-Mode-Fiber Coupling in Atmospheric Turbulence 404
Integrated Maka Antenna Transmitting and Receiving Technology 410
Automatic Spatial Optical-Coupling Alignment 413
Mode-Conversion Method 417
Adaptive-Optical Wavefront Correction 419
Correction of Non-common Path Aberrations by Reverse-Transmission Calibration 421
Comparison of Several Different Methods to Improve the Spatial Optical-Fiber Coupling Efficiency 433
Summary 435
References 436

Abstract

This chapter summarizes the research progress of spatial light to optical-fiber coupling technology in aims to improve the coupling efficiency in optical wireless communication, and introduces the research work of Xi'an University of Technology in this field, including the automatic alignment in spatial light to optical-fiber coupling, optical mode conversion in free space, adaptive optics wavefront correction in spatial light to optic fiber coupling, and design of the Maksutov–Cassegrain integrated transceiver antenna. In addition, the research progress of coupling performance by using lens array to single-mode fiber and Maksutov–Cassegrain receiving antenna in atmospheric turbulence are introduced. Finally, the development of spatial light to optical-fiber coupling technology in optical wireless communication is prospected.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_13

Keywords

Optical wireless communication · Fiber coupling · Automatic alignment · Mode conversion · Wavefront correction

Introduction

To improve its efficiency, the domestic and foreign research progress of spatial optical-fiber coupling technology in optical wireless communication is introduced, including automatic spatial optical-fiber coupling alignment technology, free-space optical mode-conversion technology, spatial optical-fiber coupling with adaptive-optical wavefront correction technology, and Maksutov–Cassegrain (Maka) antenna integrated transceiver technology.

The performance of coupling a lens array and a single lens to a single-mode fiber in atmospheric turbulence is introduced, as well as the spatial opto-coupling performance of the Maka antenna as a receiving antenna in atmospheric turbulence. Finally, the development of spatial optical-fiber coupling technology for optical wireless communication is discussed.

Foreign Research Progress

A signal light can be more easily and effectively coupled from free space to fiber with active and passive fiber-coupling components, providing efficient light field mixing for coherent-optical communications. All optical networks can transmit by coupling spatial optics into an optical fiber. This technology can simplify optical wireless communication and improve its communication rate. It is possible to control the characteristics of a signal beam in the optical fiber, promote the modularization of optical wireless communication systems, and improve system interchangeability [1], with a wide range of application prospects in military and civilian fields.

In optical wireless communication systems, the optical-fiber coupling efficiency is an important factor restricting the development of optical communication. The diameter of a single-mode fiber core is approximately 10 microns, which increases the difficulty of spatial optical coupling into optical fibers. The arrival-angle undulation under the influence of atmospheric turbulence, static-angle deviation of optical fibers, etc. will affect the coupling efficiency of optical fibers [2]. Therefore, improving the efficiency of spatial optical coupling into optical fibers and improving the system stability has become an important topic in the optical wireless communication field.

With the vigorous development of optical wireless communication, researchers at home and abroad have achieved significant research results as they have actively explored spatial optical-fiber coupling technology.

From 1990–2000, spatial light single-mode-fiber coupling technology was mainly used in the field of astronomy. In 1988, Stuart S et al. studied the coupling of starlight into a single-mode fiber. They calculated that 80% of the incident power at the pupil could be coupled to a single-mode fiber located on the central axis of the focal plane. At $D/r_0 = 4$ (D: coupling-lens diameter, r_0: atmospheric coherence length), the coupling power increased by a factor of 3.6 [3].

Since then, space optical-fiber coupling technology has been gradually applied in the field of light detection and ranging (lidar). In 1994, Wright et al. added an erbium-doped fiber amplifier (EDFA) before the fiber-coupling module of a lidar system, which increased the fiber-coupling efficiency by 20.0% [4]. In 1998, Winzer et al. applied the coupling efficiency expression of a derived monochromatic light to a fiber, after passing it through a lens and coupling it to the fiber lidar field. The numerical simulation results showed that the maximum coupling efficiency of the system was 42.0% [5].

Atmospheric turbulence has a significant impact on the efficiency of coupling spatial light to optical fibers, and different degrees of fading occur. Most scholars use adaptive optics to compensate for wavefront distortion and correct the light wave of the coupling end face to a Gaussian-distributed plane wave, thereby improving the fiber-coupling efficiency.

In 1998, Ruilier et al. studied the use of adaptive optics to improve the coupling efficiency of space-light single-mode fibers under the influence of atmospheric turbulence [6].

In 2002, Weyrauch et al. used a deformation mirror to study adaptive fiber coupling, determined the position error of the fiber by coupling the change in optical power, and combined it with the stochastic parallel gradient-descent (SPGD) algorithm to further improve the efficiency of spatial optical-fiber coupling. A maximum coupling efficiency of 60.0% was obtained in single-mode fiber coupling and 70.0% in multimode fiber coupling [7].

In 2005, Dikmelik et al. experimentally analyzed the influence of atmospheric turbulence on spatial optical-coupling efficiency and calculated that when the coupling efficiency was less than 5.0%, the communication distance could reach approximately 100 m under strong turbulence conditions. Under medium turbulence conditions, the communication distance reached approximately 800 m. Under weak turbulence conditions, the communication distance exceeded 1000 m [8].

In 2006, Toyoshima et al. analyzed the optimal relationship between the Airy spot size and fiber-mode field size on the spatial optical-fiber coupling efficiency under random angular jitter caused by atmospheric turbulence. The results showed that when the ratio of the normalized random angular jitter to the mode field radius was greater than 0.3, the average bit-error rate (BER) of the system decreased significantly [9], from approximately 10^{-1} to 10^{-4}.

In 2019, Carrizo et al. corrected the wavefront phase to increase the power coupled into a fiber by iteratively updating the phase of a single focal-plane speckle. The experimental results showed that under turbulent conditions, the power increased by 60.2% in less than 60 iterations of power measurements [10].

In satellite–ground laser communication, the influence of atmospheric turbulence on the laser-carrier signal transmission is critical. Scholars have conducted in-depth research on the principles and methods needed to overcome the influence of atmospheric turbulence on satellite–terrestrial laser communication.

Generally, the caliber of an optical system is round, and the Fran and Fei diffraction images are called Airy spots. The image of each object point is a diffuse spot, and the two diffuse spots cannot be distinguished when they are close, which limits the system resolution. The larger the spot, the lower the resolution. This is a limitation of physical optics and is caused by the diffraction of light.

In 2008, Fidler et al. studied the effect of phase distortion caused by atmospheric turbulence on the coupling efficiency of laser beams coupled to single-mode fibers in different scenarios of geostationary satellites, high-altitude platforms, and optical land stations, and proved that the performance of geostationary satellite-to-high-altitude platform communication can approach the diffraction limit when correcting the tilt component [11].

In 2010, Hiderki et al. studied the fiber-coupling efficiency loss of satellite–ground laser communication links under the influence of atmospheric turbulence. The fiber-coupling loss was more than 10 dB [12].

In 2012, Takenaka et al. studied a fast mirror that could operate at high frequencies under atmospheric turbulence conditions and verified the tracking performance of its spot position in a satellite–ground laser communication experiment. The results showed that the coupling efficiency of the satellite–ground laser-communication link faded between 10 dB and 19 dB, which was consistent with the theoretical calculation results of 17 dB [13].

In 2016, NASA conducted a satellite–ground laser experiment to test the optical systems of two ground stations. When the beam was coupled into a single-mode fiber, it was affected by atmospheric turbulence, so it was compensated by an adaptive-optical system with a double-deformation mirror design. The results showed that, under ordinary atmospheric conditions, the average coupling efficiency was higher than 50.0% [14, 15].

Because a single-mode fiber has an extremely small core and difficult optical-coupling requirements, to further improve the efficiency of spatial optical-fiber coupling, automatic coupling alignment control systems are used to adjust the position of the fiber to the best coupling position in real time to reduce the coupling efficiency loss to a certain degree.

In 1990, Lincoln Laboratory developed an active coupling method for fiber nutation using free-space optical communication. A spatial-optical single-mode fiber was automatically aligned, which reduced the coupling efficiency loss [16].

In 2001, Sayan et al. obtained a coupling efficiency of approximately 50.0% by determining the change in the incident beam angle through the feedback of a position sensor in an optical code-division multiple-access experiment with multichannel multiplexing, and adjusting a fast-tilting mirror to compensate for the error [17].

Microlens technology also plays a key role in improving the spatial optical-fiber coupling efficiency by changing the spot size and adding control algorithms. Thus, lens technology is also widely used in fiber coupling to improve the coupling efficiency.

In 1995, Modavis et al. established a Gaussian theoretical model based on a laser and a light field, analyzed the coupling performance of deformed fiber microlenses between laser diodes and single-mode fibers, and experimentally measured the average coupling efficiency of the lenses as 79.0% [18].

In 2002, Oswald et al. discussed the coupling efficiency of plane-wave coupling into a single-mode fiber with a misalignment error relative to the focal point of the lens. Studies showed that a collimation system reduced the error and the coupling efficiency reached 61.0% [19]. In the same year, Sherman et al. designed a radio optical communication system that could position an $n \times n$ optical array on the focal plane of a Ritchey–Chretien (RC) telescope with an aspheric mirror. This enabled point-to-multipoint communication with a single optical telescope [20].

In 2010, Danuel et al. designed a method for spatial light coupling using an optical-fiber array. Piezoelectric ceramics were used to drive the microlens position to adopt a control algorithm with the maximum output energy. The experimental results showed that the output optical power increased by 39 dB, and the coupling efficiency of the optical fiber was improved [21].

In 2018, Hottinger designed a single-mode fiber-coupled angle sensor based on a microlens. It detected the change in light energy in the fiber, determined the fiber position error, and improved the coupling efficiency [22].

In recent years, scholars have extensively researched the coupling efficiency of multimode fibers, low-mode fibers, and specially designed fibers to improve the low coupling efficiency of space beams caused by small single-mode fiber cores.

In 1997, Keming et al. investigated a technique for coupling the output of a high-power diode laser into a multimode fiber that included two microlenses for beam shaping. Experiments showed that the overall coupling efficiency of the diode laser to fiber was 71.0% [23].

In 2007, Horton et al. investigated the effect of diffraction on the coupling efficiency of single-mode and low-mode fibers. The results showed that the low-mode fiber had a higher maximum coupling efficiency (>90.0%) [24]. In 2013, NASA conducted a lunar–Earth laser communication demonstration experiment, in which a multimode optical fiber was used to couple spatial light. The results showed that the coupling efficiency of the space-light multimode fiber under atmospheric turbulence could reach 92.0% [25, 26].

Because the signal light has different degrees of efficiency loss in the process of coupling into a single-mode fiber after transmission through a multimode fiber, Carl M et al. proposed a new device that improved free-space light coupling to a single-mode fiber in 2016. It consisted of an optical-fiber cone that acted as a mode filter to filter out high-order modes and then coupled other modes into the single-mode fiber, resulting in a power loss of 42.0% [27].

The relevant foreign research progress is presented in Table 13.1.

Table 13.1 Foreign research progress

Year	Person/Organization	Research progress
1988	Shaklan/National Optical Astronomy Observatory	Starlight was coupled into a single-mode fiber
1990	Boroson/MIT Lincoln Laboratory	Active coupling scheme based on fiber nutation for laser communication
1995	Modavis/Corning, Inc.	Established theoretical models of scalar diffraction and Gaussian pattern shapes of laser and optical-fiber fields
1998	Winzer/Institut fur Nachrichtentechnik und Hochfrequenztechnik	Optimized the receiving aperture size to improve the coupling efficiency
2002	Oswald/Vienna University of Technology	Collimation system increased the coupling efficiency
2002	Weyrauch/US Army Research Laboratory	Studied adaptive fiber coupling with a deformable mirror
2002	Smolyaninov/University of Maryland	EDFA was coupled to a single-mode fiber as a receiving preamplifier
2005	Dikmelik/Johns Hopkins University	Coherent fiber arrays improved the fiber-coupling efficiency
2006	Toyoshima/Johns Hopkins University	Coupling efficiency of space light caused by atmospheric turbulence
2007	Horton/MIT Lincoln Laboratory	Effect of diffraction on coupling efficiency of single-mode and low-mode fibers
2010	Hiderki/University of Electro-Communications	Maximum fiber-coupling efficiency in the presence of random angular jitter
2010	Danuel/Johns Hopkins University Applied Physics Laboratory	Added an algorithm to a microlens to improve coupling efficiency
2012	Takenaka/Space Communication System Laboratory	Jitter error was compensated to improve fiber-coupling efficiency
2013	NASA	Lunar–Earth laser communication demonstration experiment
2014	Carl/Heinrich Hertz Telecommunications Institute	Conical fiber added between multimode and single-mode fibers to improve coupling efficiency
2016	NASA	Adaptive-optical system was used to compensate atmospheric turbulence and the beam was coupled into a single-mode fiber
2018	Hottinger/Astronomie der Universitat Heidelberg	Angle sensor was used to align spatial optical-fiber coupling
2019	Carrizo/Institute of Communication and Navigation	Phase of each focal-plane speckle was iteratively updated to correct the wavefront and improve the coupling efficiency

Domestic Research Progress

Domestic scholars have also carried out in-depth research and made great progress in space optical-fiber coupling technology. Owing to the complexity of space optical-channel transmissions, the intensity of atmospheric turbulence affects the coupling efficiency of space optical fibers.

In 2006, Xiang et al. studied the efficiency of spatial-light coupling to a single-mode fiber and analyzed the change in optical power under the influences of multiple factors. They found that the uplink had little effect on the optical coupling efficiency. For the downlink, the relative power decreased when it was affected by the aperture average effect and increased when it was affected by turbulence [28].

In 2009, Ma et al. studied the influence of arrival-angle fluctuation on fiber-coupling efficiency and obtained the optimal design parameter value of a fiber-coupling optical system with random jitter. When the ratio of the aperture radius to the fiber-mode field radius $\beta = 1.121$, a maximum coupling efficiency of 81.0% was obtained [29].

In 2018, Liu et al. studied the effect of fiber-alignment error and random angle jitter on fiber-coupling efficiency and obtained a maximum coupling efficiency of 62.0% [30].

In 2021, Song et al. established a model to calculate the transient coupling efficiency. They obtained the coupling efficiency change rule with the ratio of the aperture radius to atmospheric coherence length, as well as the change in geometric coupling parameters [31].

Owing to the phase distortion of the beam wavefront caused by atmospheric turbulence, the coupling efficiency of a space optical fiber is significantly reduced, which seriously affects the communication quality. To suppress the influence of atmospheric turbulence on the coupling efficiency of space optical fibers, scholars have used adaptive-optical systems to further improve the coupling efficiency.

In 2010, Wu et al. proposed a method that combined adaptive-optics technology with coherent fiber array technology to improve the efficiency of spatial optical coupling. The results showed that this method improved the efficiency of signal light coupling into optical fibers after atmospheric turbulence from 3.0% to 39.0% [32].

In 2011, Han et al. proposed a model-free blind-optimized wavefront correction technique to improve the spatial optical-coupling efficiency under the influence of atmospheric turbulence. Their results showed that the coupling efficiency of the single-mode fiber improved from 6.0% to approximately 60.0% [33].

In 2012, Yang et al. used the wavefront phase mode to compensate for a beam affected by atmospheric turbulence and coupled it into a single-mode fiber, achieving a coupling efficiency of approximately 81.0% [34].

In 2013, Xiong et al. used a 37-unit adaptive-optical system to compensate for atmospheric turbulence aberrations. Studies showed that the coupling efficiency increased by 20.0% when the Strehl ratio increased from 0.16–0.35 [35].

In 2014, Luo et al. found that the fiber-coupling efficiency was affected by the wave difference caused by a single aberration. The average coupling efficiency

increased from 30.1–61.7% after using an adaptive fiber-optic source collimator for compensation [36].

In 2015, Li Feng team developed an adaptive optical-fiber coupler and combined it with the SPGD algorithm to select several systems under different atmospheric turbulence conditions for closed-loop control. The results showed that when optimal SPGD control parameters were selected, the fiber-coupling efficiency was significantly improved, from 40.0% before convergence to 76.0% after convergence [37].

Atmospheric turbulence not only distorts the beam wavefront phase but also degrades the beam mode, resulting in an incomplete match with the receiving single-mode-fiber mode field. The energy distribution of a beam can be concentrated by means of a mode-conversion method, and the converted fundamental mode form can be collected by a single-mode fiber, to improve the coupling efficiency. In 2015, Qi et al. used a spatial-light modulator (SLM) to convert LP01 mode into a high-order mode. This method was simple, easy to implement, had low device requirements, and high repeatability [38].

In 2017, Tu et al. designed a multiplexer/demultiplexer based on a three-core fiber and supported LP01, LP11a, and LP11b spatial-mode transmissions. The power conversion between the LP11a and LP01 modes and between the LP11b and LP01 modes was simulated and analyzed, and the mode-conversion efficiency was 90.0% [39].

In 2018, Shin et al. proposed a new optical waveguide LP01–LP02 mode converter using a combination of two pyramidal structures. The conversion efficiency of this mode converter was 90.0%, which greatly reduced the crosstalk with other modes [40].

The coupling process is demanding and difficult, owing to the small core of a single-mode optical fiber. Research on automatic spatial optical-fiber coupling alignment technology has solved this problem to a great extent. In 2007, Gao et al. completed automatic coupling through a closed-loop control system they designed. Experiments showed that this method significantly reduced the automatic alignment time and achieved a maximum coupling efficiency of 510.2% [41].

In 2017, Wu et al. proposed an SPGD algorithm and a scheme based on a helical-grating scanning algorithm to improve the spatial light-fiber coupling efficiency. The theoretical simulation results showed that the spatial light-fiber coupling efficiency was up to 81.0% [42].

In 2019, Baiqiu Zhao et al. combined two techniques of chapter-dynamic coupling algorithm and fast reflector of fiber optic photodetector together to obtain a maximum coupling efficiency of 510.6% under nondynamic conditions [43].

In 2019, Li et al. proposed an adaptive-coupling method for optical communication systems based on thick–fine laser nutation technology. The experimental results showed that, under simulated atmospheric turbulence, the average coupling efficiency of the system was improved to 62.4% [44].

In 2019, Qi et al. applied a pattern-search method to spatial optical-fiber coupling to achieve alignment coupling in X, Y, and Z degrees of freedom. This shortened the alignment time to within 25 s and achieved a success rate of over 90.0% [45].

Research on improving the spatial light-fiber coupling efficiency not only involves the receiving end, but also the light source. For example, the fiber-coupling output mode of a semiconductor laser can improve the asymmetry of the output beam and realize optical coupling between multiple semiconductor light sources, thus improving the output power and fiber-coupling efficiency.

In 2000, Shi et al. proposed a beam-shaping technology for high-power laser-array devices that achieved a coupling efficiency of 50.0% [46].

In 2001, Bo et al. used a cylindrical lens to collect and collimate the output beam of an array semiconductor laser and finally coupled the output beam to a multimode fiber. The coupling efficiency was as high as 75.0% [47].

In 2004, Zhou et al. calculated the surface-shape formula of two groups of orthogonal collimating microlens arrays and conducted coupling experiments between high-power semiconductor lasers and fibers. They obtained a coupling efficiency greater than 53.0% [48].

In 2010, Wang et al. developed a fiber-coupling module with excellent performance using a step-mirror group for beam shaping a semiconductor laser array, and the coupling efficiency reached 73.8% [49].

In 2014, Chen et al. coupled four 980-nm single-tube semiconductor lasers into a fiber after combining the beams and obtained a coupling efficiency of 51.7% [50].

In 2015, Xu et al. developed an efficient and high-power array of fiber-coupled semiconductor-laser modules using beam-shaping and spatial-beam combination methods. The experimental results showed that the optical output power of the fiber could reach 327 W, and the coupling efficiency was greater than 90.0% [51].

In 2015, Liu et al. designed a multibeam fiber-coupling system. After coupling beams into a fiber through a ZEMAX simulation, the coupling efficiency exceeded 98.0% [52].

In 2017, Liu et al. successfully developed a high-power semiconductor-laser fiber coupling module using polarization and space-coupling technology. The experimental results showed that the fiber-coupling efficiency was 60.0% [53].

An optical-fiber lens can change the size and shape of the optical-fiber mode field. Microlenses have the advantages of changing the spot size and adding a control algorithm; therefore, optical-fiber lenses or lens technology is also widely used in optical-fiber coupling to improve the coupling efficiency.

In 2002, Wang et al. used a multimode fiber with a small aperture to replace the microlens and coupled the output beam of a laser-diode line array into a multimode fiber array. The coupling efficiency measured in the experiment was 75.0% [54].

In 2002, Lu et al. studied a coupling method of placing fiber microlenses between semiconductor lasers and fibers, and experimentally measured coupling efficiencies up to 80.0% [55].

In 2003, Wei et al. proposed the use of a telescope collimator system to improve the fiber-coupling efficiency. The experimentally measured coupling efficiencies of a polarization-maintaining fiber and single-mode fiber were 68.0% and 70.0%, respectively [56].

In 2003, Zhao et al. established a theoretical model of the coupling between a cone-end spherical-lens fiber and a flat-end fiber. The results showed that the

maximum coupling efficiency was 79.3% when the spherical radius was 25.6 μm and cone angle was 110 ° [57].

In 2014, Liu et al. processed fiber end faces into wedge-shaped microlenses and found that the coupling between a semiconductor laser and fiber could satisfy phase and mode field matching, with a maximum coupling efficiency of 81.3% [58].

In 2016, Shi et al. designed cylindrical and aspherical lenses using geometric optics and verified the designed lens system using the ZEMAX simulation software. The simulation results showed that the coupling efficiency was improved by approximately 12.0% [59].

Optimizing the self-coupling system and fiber structure makes the system easier to process and design, and the anti-interference ability is stronger, which improves the fiber-coupling efficiency. In 1981, Li et al. coupled a light-emitting tube with a ball-end fiber, and obtained a maximum coupling efficiency of 14.0% in the final experiment [60].

In 2008, Feng et al. nutated the coupling of a Gaussian beam to a multimode fiber using a novel optical-antenna communication system. Theoretical calculation showed that when the laser wavelengths were 1310 nm and 1550 nm, the maximum coupling efficiencies were 81.5% and 82.5%, respectively [61].

In 2011, Wang et al. used a specially designed miniaturized optical antenna to reduce the impact of turbulent atmosphere on the communication quality. The results showed that the receiving sensitivity of this system was greater than −35 dBm [62] at a certain communication distance and rate.

In 2013, Hu et al. established a conical multimode fiber transmission model, which obtained a coupling efficiency of nearly 80.0% through theoretical calculation [63].

In 2014, Zhang et al. welded single-mode fibers with large and small cores, and then coupled a spatial beam from one side of the large core. Experiments showed that this method could achieve a fiber-coupling efficiency of 60.0% [64].

In 2019, Yan et al. proposed using conical fiber in a optical wireless communication coupling system, which improved the lateral-migration capacity difference and increased the coupling efficiency [65].

In 2021, Ren et al. designed a structure with a symmetrical front flange that was suitable for a high-temperature environment and had a high coupling efficiency. The simulation showed that when the temperature field was 40 °C, the coupling efficiency loss of the front-flange symmetric structure was reduced by approximately 61.0%, compared with the bottom-mounted structure [66].

In a spatial optical-fiber coupling system, in addition to external factors affecting the fiber-coupling efficiency, incident light waves and fiber end faces have many static errors and affect the space optical-fiber coupling efficiency.

In 2013, Zhang et al. found that the central axis of the coupling lens deviated from the single-mode fiber, and the incident light-wave mode was poorly matched with the single-mode fiber. This led to a decrease in the coupling efficiency of the single-

mode fiber. Through simulation, it was found that the coupling efficiency was influenced more by tilt and lateral deviations than by defocusing [67].

In 2013, Luo et al. found that the coupling efficiency was greater than 80.0% when the off-axis value between the multimode fiber and the outbound beam of a Cassegrain antenna was less than 4.64 μm and the antenna off-axis angle was less than 0.055 rad. When the off-axis angle or axial quantity increased, the coupling efficiency decreased [68].

The relevant research progress in China is presented in Table 13.2.

Table 13.2 Domestic research progress

Year	Person/Organization	Research progress
2000	Shi Peng/Chinese Academy of Sciences	High-power laser-diode linear-array device beams can be shaped by a microwafer prism stack
2002 2003	Wang Xiaowei/Chinese Academy of Sciences Wei Rong/Chinese Academy of Sciences	Output beam of a laser-diode linear array is coupled to a multimode fiber array A telescope collimation system is used to improve the laser-fiber coupling efficiency
2003	Zhao Faying/Zhejiang University	Theoretical coupling model between a flat-end fiber and cone-end spherical lens fiber
2004	Zhou Chongxi/Chinese Academy of Sciences	High-power semiconductor laser-array beam improves the coupling efficiency
2009	Ma Jing/Harbin Institute of Technology	Effect of the arrival-angle fluctuation on fiber-coupling efficiency
2010	Wu Hanling/University of Slaughter	Hybrid technology of an adaptive-optical system and coherent fiber array technology
2011	Han Liqiang/Harbin Institute of Technology	Model-free blind optimization of wavefront correction improves coupling efficiency
2012	Yang Qingbo/Harbin Institute of Technology	The wavefront phase mode is used to compensate a beam passing through atmospheric turbulence
2013	Luo Zhihua/University of Electronic Science and Technology of China	Coupling efficiency of a Cassegrain system for collimation and deflection
2014	Luo Wen/Chinese Academy of Sciences	Coupling efficiency affected by the wave difference caused by a single aberration
2014	Liu Yangyang/Hebei University of Technology	Coupling of a single-mode fiber and semiconductor laser
2018	Liu Yutong/Changchun University of Science and Technology	Effect of random angle jitter and fiber-alignment errors on coupling efficiency
2019	Zhao Baiqiu/Changchun University of Science and Technology	Coupling efficiency is improved by a combination of a fast mirror and fiber laser nutation algorithm
2021	Ren Lanxu/China Academy of Space Technology	Symmetrical structure of the front flange improves coupling efficiency

Research Progress on Spatial Optical-Fiber Coupling Technology

Since 2000, Professor Ke Xizheng's team from Xi'an University of Technology has conducted extensive research on the optical design of communication systems [69], coding-modulation technology for signal transmissions [70], optical wireless coherence detection technology [71, 72], adaptive-optics technology [73], etc. The team also successfully developed a differential phase-shift keying (DPSK)-modulated coherent detection system terminal and constructed communication links with 600-m, 1.3-km, 10-km, and 100-km distances for high-definition video transmissions.

To effectively improve the fiber-coupling efficiency and communication quality, Professor Ke's team has systematically investigated the field of space optical-fiber coupling technology in optical wireless communication. Chen [74] analyzed the multimode fiber coupling problem under the influence of weak turbulence. The results showed that the coupling efficiency of multimode fibers decreased less than that of single-mode fibers under the influence of random spot jitter.

Lei [75] proposed to use the off-axis Gaussian-beam emission of a Maksutov–Cassegrain (Maka) antenna array and spatial light coupling technology to integrate the antenna's transmitter and receiver. The experimental results showed that the transmission efficiency of the Maka antenna was significantly improved. Luo [2] designed an automatic-alignment closed-loop device system to complete the coupling. The system used piezoelectric ceramics to control the optical fibers. Experiments showed that the system improved the alignment, positioning time, and the coupling efficiency of the optical fiber.

Zhang [76] found that mode conversions greatly improved the fiber-coupling efficiency. Yang [77] studied the influence of wavefront distortion at different distances on the coupling efficiency. The experimental results showed that the lifting effect of long-distance correction was inferior to that of short-range correction and that the numerical fluctuation of long-distance coupling efficiency was greater than that of short-range coupling efficiency.

Yin [78, 79] designed a single-mode fiber-coupling device system based on piezoelectric ceramics. The experiment automatically aligned a space light and single-mode fiber through the control algorithm, and the coupling efficiency was increased to 53.2%. Wu [80] designed a fiber coupler with a five–degrees of freedom (5-DOF) coupling device. The experimental results showed that the 5-DOF fiber coupler corrected different alignment errors, and the single-mode fiber coupling efficiency of the system was significantly improved after the closed loop.

Space-Light–Single-Mode-Fiber Coupling in Atmospheric Turbulence

Lei et al. compared the single-mode fiber coupling efficiency of a lens array with that of a single lens under atmospheric turbulence conditions and analyzed the spatial optical-coupling performance of a Maka antenna as a receiving antenna.

Coupling Efficiency of a Lens Array in Atmospheric Turbulence

Figure 13.1 shows planar and physical drawings of a lens and lens array of equal area. Figure 13.1a shows a single lens with a diameter of D_A, and Fig. 13.1b shows a lens array with a total diameter of D_A, consisting of an unit lens array with a diameter of $d_A = D_A/3$. It can be seen from the figure that the total areas of the single lens and lens array have the same relative aperture.

Figure 13.2 shows the coupling efficiency curve of the equal-area lens and lens array under medium–strong turbulence, where a is the coupling parameter, C_n^2 is the structural constant of the atmospheric refractive index, l_0 is the internal turbulence

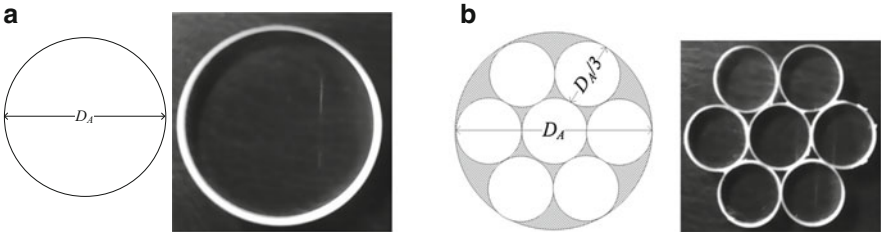


Fig. 13.1 Planar diagram of equal-area lens and lens array [76]. (**a**) Lens. (**b**) Lens array

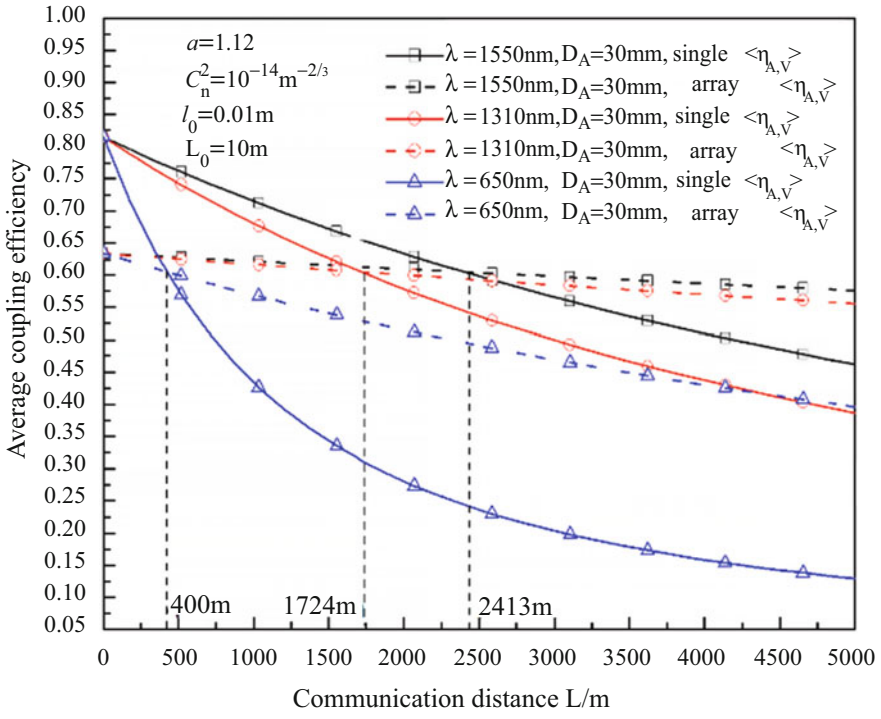


Fig. 13.2 Coupling efficiency curve of a lens and lens array under medium–strong turbulence [75]

scale, L_0 is the outer turbulence scale, and λ is the wavelength. It can be seen that the coupling efficiency of the single lens is higher than that of the lens array in short-distance communication. This is because the gaps between the lens units in the lens array are not fully utilized; hence, part of the beam cannot be fully coupled into the fiber through the gaps, resulting in a certain degree of coupling efficiency loss [81].

When the communication distance increased gradually, the coupling efficiency of the lens array decreased slowly with an increase in the communication distance, owing to the turbulence intensity. When the transmission distance exceeded a certain value, the coupling efficiency of the lens array was higher than that of a single lens with an equal area. This is because the diameter of an array element is smaller than that of the single lens; therefore, the number of speckles on the surface of the lens array is much smaller than that of the single lens, resulting in an increase in the coupling efficiency [82].

Figure 13.3 shows the variation of the power-lift coefficient K of the lens array relative to the single lens with the communication distance under medium–strong turbulence. When the communication wavelengths are 650, 1310, and 1550 nm and the communication distances are greater than 400, 1724, and 2413 m, respectively,

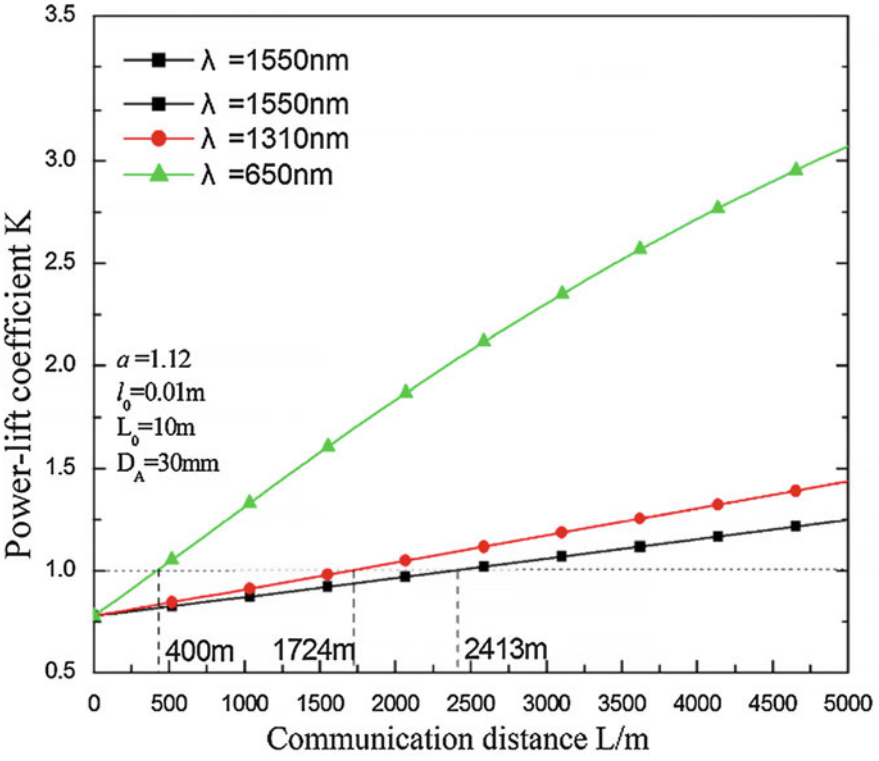


Fig. 13.3 Power-lift coefficient K of the array relative to the lens under medium–strong turbulence with varying communication distance [75]

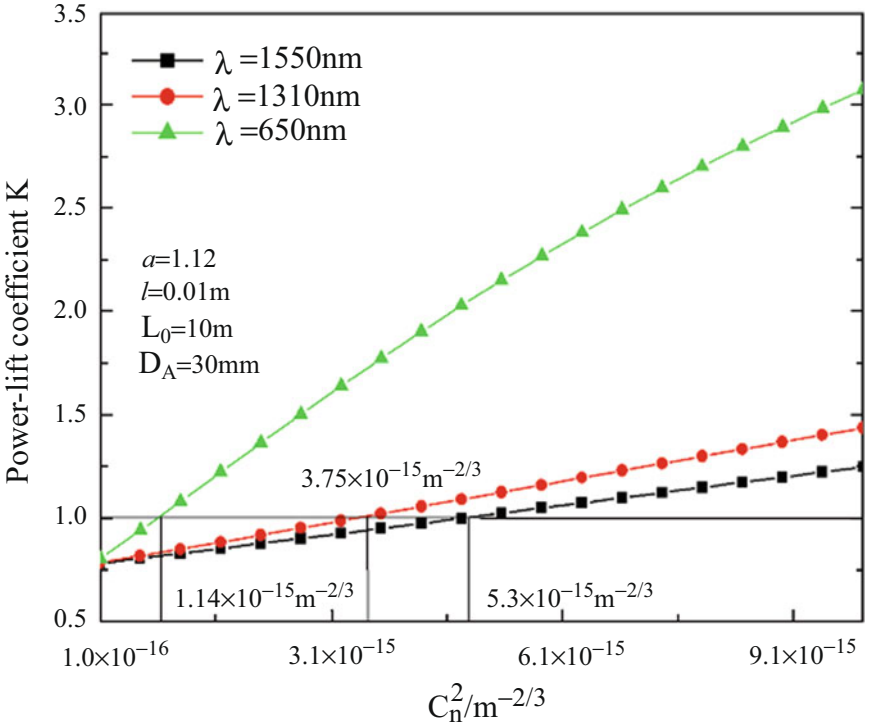


Fig. 13.4 Power-lift coefficient K of the array relative to the lens with a fixed communication distance (5 km) and varying turbulence intensity [75]

the optical power of the coupled lens array exceeds that of the coupled single lens, and a conclusion consistent with Fig. 13.1 is drawn.

Figure 13.4 shows that the power-lift coefficient K of the lens array relative to the single lens varies with the turbulence intensity when the communication distance is 5 km. When different wavelengths exceed a certain turbulence intensity, the optical power of the coupled lens array exceeds that of the coupled single lens.

Coupling Efficiency of a Maka Antenna Under Atmospheric Turbulence

Figure 13.5 shows a schematic diagram of the structure of a retractable Maka antenna. This structure makes the Maka antenna small, lightweight, and easy to fabricate. When it is used as a spatial optical-coupling system, it enables large-aperture coupling for the receiver and high-light power convergence technology.

Lei et al. studied the coupling efficiency of spatial fiber when a Maka antenna was considered equivalent to a lens system and used as a receiving antenna under the influence of atmospheric turbulence.

Figure 13.6 shows the curves of coupling parameter a and the coupling efficiency under different shielding ratios b with weak turbulence ($C_n^2 = 10^{-16} m^{-2/3}$) for a horizontal transmission. L is the communication distance, λ is the wavelength, D is

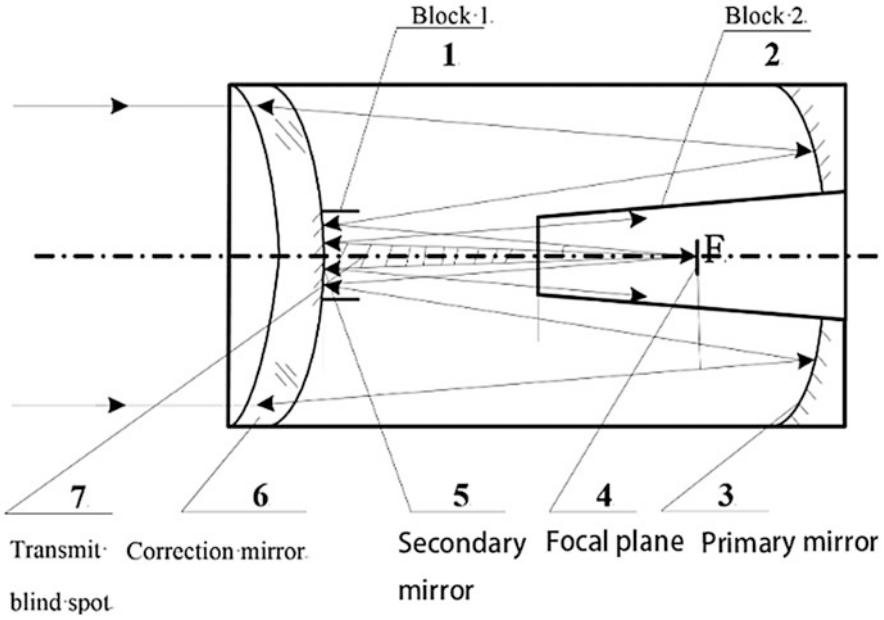


Fig. 13.5 Structural diagram of a retractable Maka antenna [75]

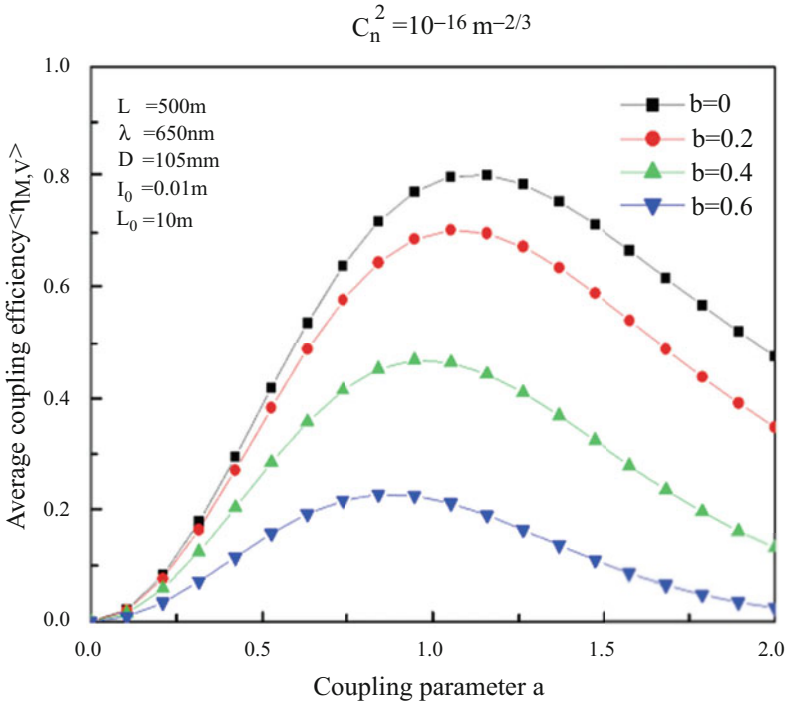


Fig. 13.6 Coupling parameters and coupling efficiency in weak turbulence [74]

the diameter of the Maka antenna, l_0 is the inner turbulence scale, and L_0 is the outer turbulence scale. It can be observed that when the shielding ratio is constant, the coupling efficiency first increases and then decreases with an increase in the coupling parameters. When the coupling parameters are fixed, the maximum coupling efficiency increases with an increase in the shielding ratio.

Figure 13.7 shows how the coupling efficiency varies with the communication distance for a horizontal transmission and weak turbulence ($C_n^2 = 10^{-16}m^{-2/3}$). With an increase in the communication distance, the coupling efficiency decreases slowly under different wavelength conditions. The coupling efficiency when the shielding ratio $b = 0$ is significantly greater than that when the shielding ratio $b = 2$. Therefore, in the case of weak turbulence, the most important factor affecting the change in coupling efficiency is the shielding ratio of the Maka antenna.

Figure 13.8 shows the curve of the coupling efficiency changing with the coupling parameters under different shielding ratios b for a horizontal transmission with medium–strong turbulence ($C_n^2 = 10^{-14}m^{-2/3}$). It can be observed that when the shielding ratio is constant, the coupling efficiency first increases and then decreases with an increase in the coupling parameters. The smaller the shielding ratio, the larger the maximum coupling efficiency. When the coupling parameters are fixed, the maximum coupling efficiency decreases rapidly with the increase of the shielding ratio.

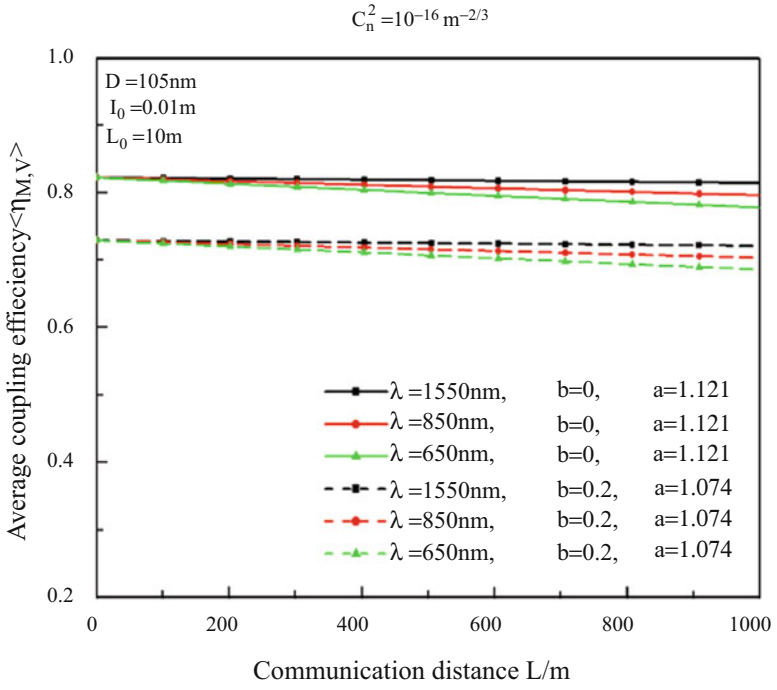


Fig. 13.7 Communication distance and coupling efficiency curve in weak turbulence [75]

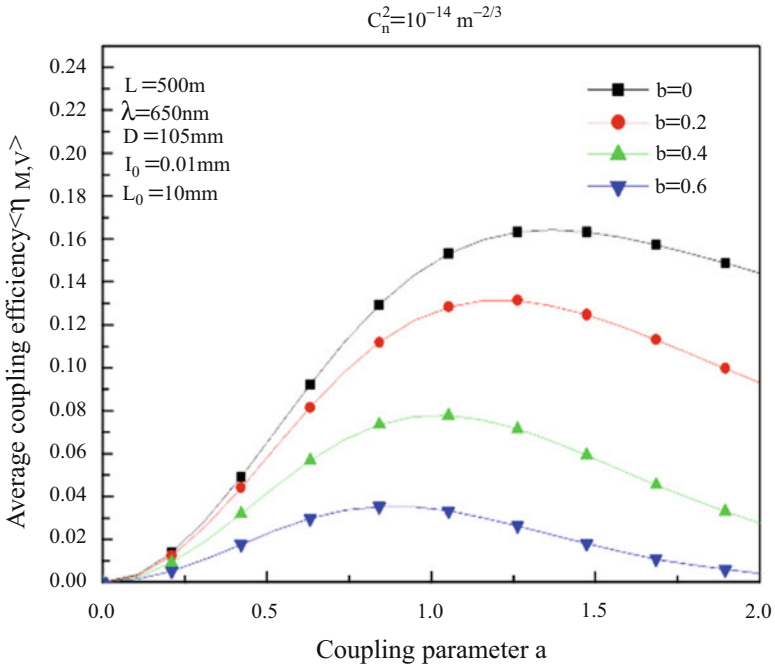


Fig. 13.8 Coupling parameters and coupling efficiency in medium–strong turbulence [75]

Figure 13.9 shows the curve of the coupling efficiency changing with the communication distance for a horizontal transmission with medium–strong turbulence ($C_n^2 = 10^{-14}m^{-2/3}$). It can be seen that with an increase in the communication distance, the longer the wavelength, the smaller the decrease in coupling efficiency. In the case of medium–strong turbulence, the most important factor affecting the coupling efficiency is the transmission distance.

Comparison of the Coupling System with a Lens Array and a Maka Antenna Under Atmospheric Turbulence

Table 13.3 shows the advantages and disadvantages of the coupling system with a single lens, lens array, and Maka antenna under atmospheric turbulence, as well as the research results of coupling efficiency changes under various influencing factors.

Integrated Maka Antenna Transmitting and Receiving Technology

Figure 13.10 shows a schematic diagram of the Maka antenna integrated transceiver experiment. The effective receiving area of this structure is the blind transmitting area of the cassette antenna. From the perspective of the transmitting terminal, the

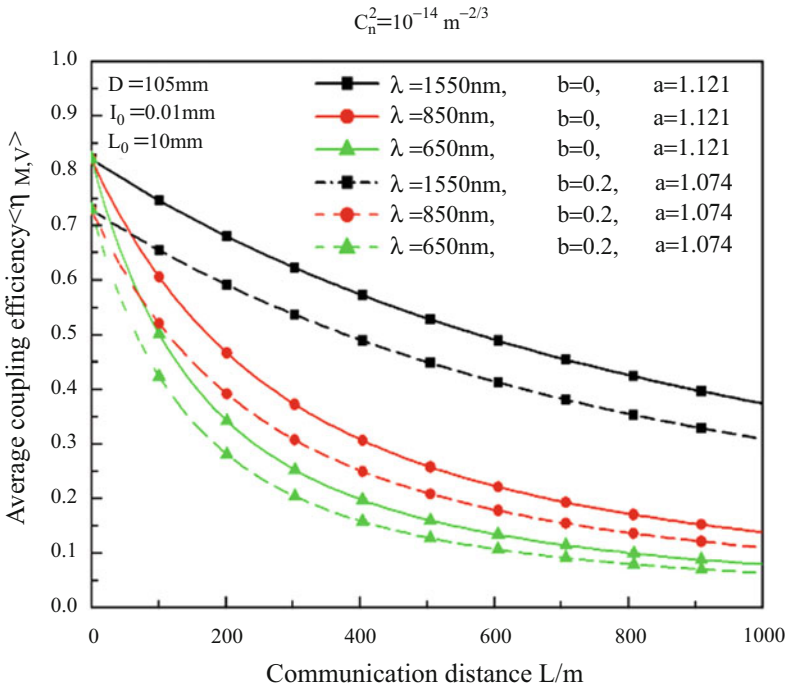


Fig. 13.9 Communication distance and coupling efficiency curve in medium–strong turbulence [75]

effective transmitting area can be maximized by changing the beam-waist size and divergence angle.

The experimental results show that the output light power of the collimator is 16.8 mW, and that of Maka antenna A is 14.9 mW. The calculated transmission efficiency is approximately 94.0%. Compared to the previously used axial-transmission method, the transmission efficiency of this method is increased by 85.0%, and the optical-power loss caused by secondary-mirror occlusion was successfully solved. The experimental results are listed in Table 13.4.

As can be seen from Table 13.4, when the distance between Maka antennas A and B was 5 m or 15 m, the received optical power fluctuated at approximately 3 mW. This is because the distance between the two antennas is relatively small, so the received light spots are generally located in the effective receiving area of the antenna. Hence, there is no significant difference in the received optical power. When the distance between the two antennas is 30 m, the transmitted beam diverges, and only part of the light spot is located in the effective receiving area, resulting in a partial power loss.

Table 13.3 Comparison of coupling systems of a single lens, lens array, and Maka antenna under atmospheric turbulence

Lens	Advantages	Shortcomings	Coupling efficiency variation
Equal-area lens (compared to array)	Relatively large effective area and all beams incident in the aperture plane can be coupled into the optical fiber	Weak ability to inhibit turbulence	Coupling efficiency of an equal-area lens is higher than that of a lens array for shorter communication distances
Lens array	Better ability to restrain turbulence in turbulent flow	The effective area is 0.778, and all the beams incident in the aperture plane cannot be coupled into the fiber	Coupling efficiency of a lens array is higher than that of an equal-area lens when the communication distance and turbulence intensity increase
Maka antenna	Large-aperture coupling receiving technology and strong convergent optical-power capability	Secondary-mirror occlusion causes energy loss	In weak turbulence, the blocking ratio of the Macca antenna is the main factor that reduces the coupling efficiency; in strong turbulence, the coupling efficiency of short wavelengths decreases to less than 20% when the transmission distance exceeds the order of km

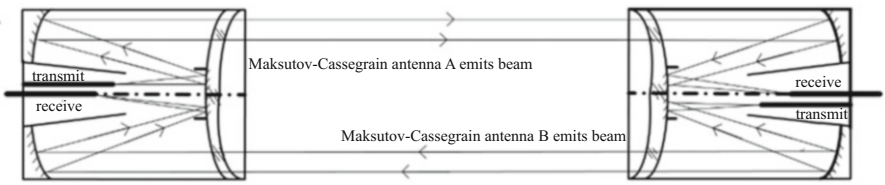


Fig. 13.10 Schematic diagram of the integrated transceiver experiment [74]

Table 13.4 Maka antenna integrated transceiver experimental data [75]

Distance/m	Antenna B emits optical power/mW	Antenna A receives optical power/mW	Coupling efficiency of Antenna A (%)	Antenna A emits optical power/mW	Antenna B receives optical power/mW	Coupling efficiency of Antenna B (%)
5		3.1	21.6		2.8	20.2
15	14.3	3.0	20.9	14.8	2.9	20.9
30		2.3	16.1		2.2	15.8

Automatic Spatial Optical-Coupling Alignment

Automatic spatial optical-coupling alignment has a direct influence on the coupling efficiency. Luo et al. studied the simulated-annealing algorithm [2] and stochastic parallel gradient-descent (SPGD) algorithm to control the voltage on piezoelectric ceramics, drive the fiber to generate a displacement, adjust the fiber position in real time, align the optical axis, and improve the coupling efficiency.

Application of a Simulated-Annealing Algorithm to Automatic Spatial Optical-Coupling Alignment

To improve the efficiency of coupling spatial light to an optical fiber, the voltage-signal output by the photodetector is used as a feedback signal. The voltage on the two-dimensional piezoelectric ceramics is adjusted by a simulated-annealing algorithm to dynamically adjust the optical-axis alignment. Figure 13.11 shows the automatic space optical-coupling alignment system using two-dimensional piezoelectric ceramics.

Figure 13.12a shows the voltage of the X-axis output after the simulated-annealing algorithm was adopted. As the number of iterations increases in the initial

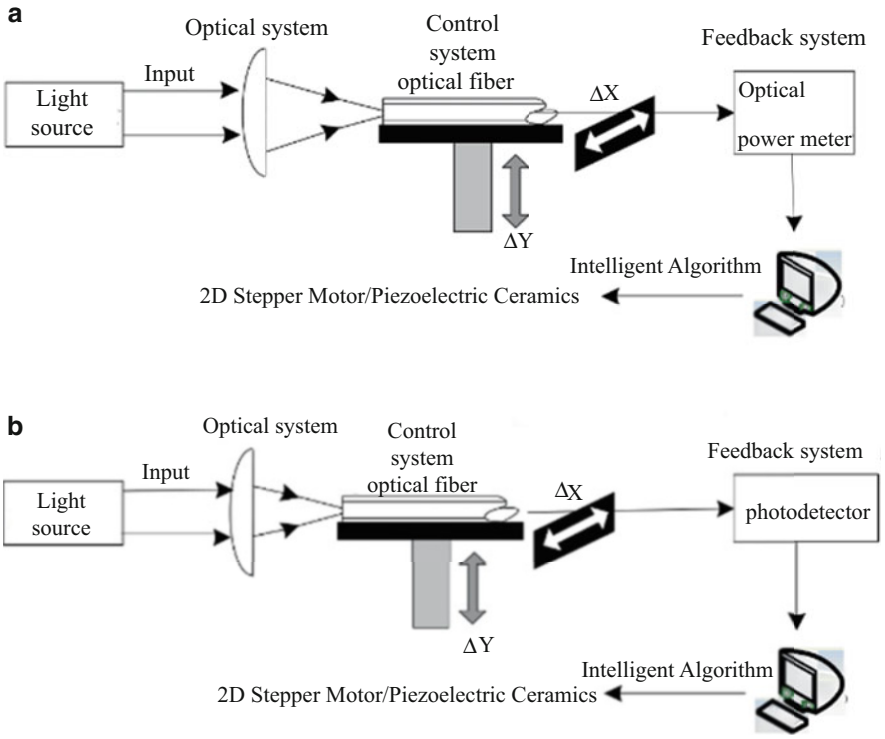


Fig. 13.11 Automatic space optical-fiber coupling alignment system [83]. (**a**) Optical signal acquisition system. (**b**) Electrical signal acquisition system

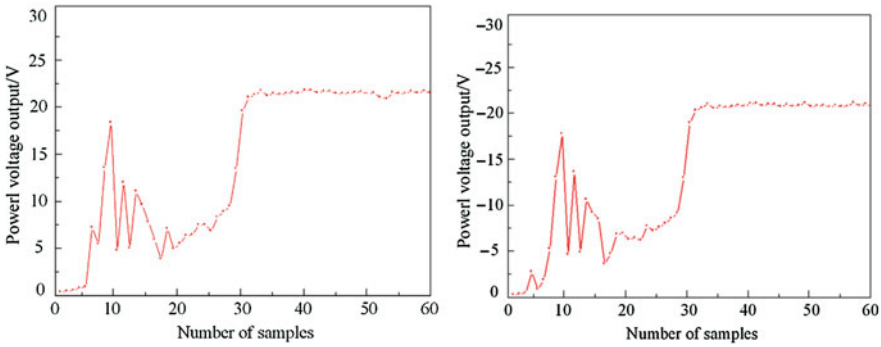


Fig. 13.12 Voltage variation when the simulated-annealing algorithm is added [2]. (**a**) *X* axis. (**b**) *Y* axis

stage, the curve fluctuates significantly. The curve fluctuation reflects the process of the system searching for the optimal coupling position. When the number of iterations is greater than 30, the curve stabilizes because the system initially searches for the optimal voltage. The voltage fluctuates slightly around 20.68 V and stabilizes, indicating that the optimal coupling position has been determined by the system.

Figure 13.12b shows the Y-axis output voltage after the simulated-annealing algorithm is used. It can be seen that the curve fluctuates with an increase in the number of iterations in the initial stage. After the fluctuation, the voltage value fluctuates slightly around −20.323 V and stabilizes; that is, the optimal coupling position has been determined. The voltage graph reflects the entire search process.

The simulated-annealing algorithm can be used to determine the optimal position in a short time, and the coupling efficiency was as high as 51.4%. The self-designed alignment system has the advantages of a reliable fixed mode and high-precision driving circuit. This is a highly efficient method for improving the precision alignment of optical fibers.

Application of the SPGD Algorithm to Automatic Spatial Optical-Coupling Alignment

To achieve higher-precision regulation of the fiber end face, Yin et al. designed a constant-voltage DC power supply and voltage-amplifying circuit corresponding to the three working conditions of piezoelectric ceramics by using three different displacement modes. Using the SPGD algorithm, the optical power in the coupled single-mode fiber is taken as real-time feedback to automatically align the coupling of the spatial optical and single-mode fibers, so that the end face of the fiber is in the best coupling position. Figure 13.13 shows the automatic-alignment block diagram of the five-dimensional piezoelectric ceramics in the space optical coupling.

Figure 13.14 shows the curve of the coupling optical power changing with the number of iterations under the influence of axial deviation. It can be observed that the coupling power of the single-mode optical fiber increases with an increase in the

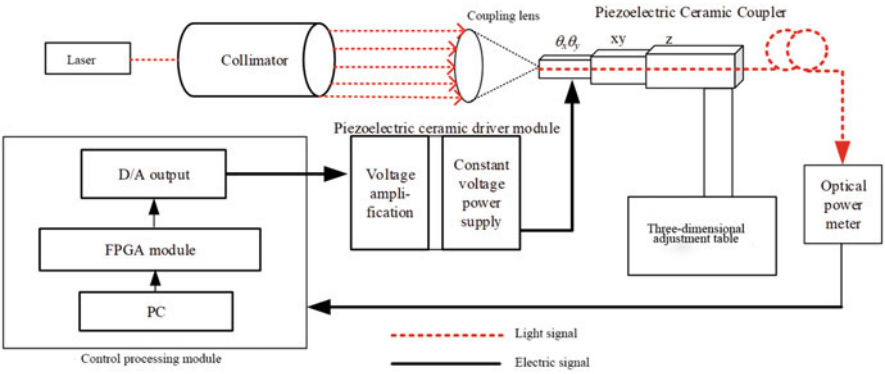


Fig. 13.13 Schematic diagram of the automatic spatial optical-coupling alignment experiment [78]

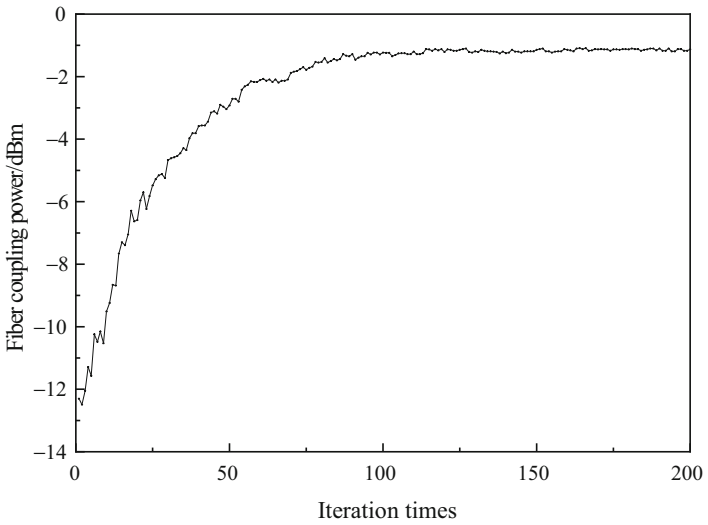


Fig. 13.14 Coupling optical-power variation with axial deviation with a varying number of iterations [78]

number of algorithm iterations, which means that the axial deviation is effectively compensated.

Figure 13.15 shows the curve of the coupling optical-power changing with the number of iterations under the influence of radial deviation. It can be observed that the coupling power of the single-mode optical fiber increases with an increase in the number of iterations, which means that the radial deviation is effectively compensated.

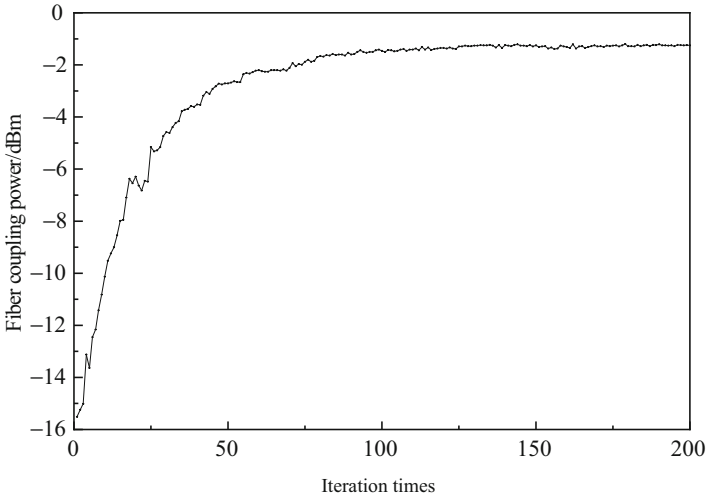


Fig. 13.15 Coupling optical-power variation with radial deviation with a varying number of iterations [78]

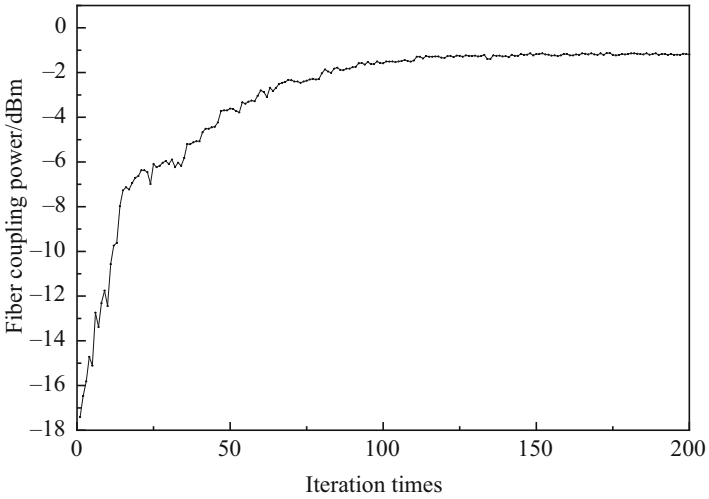


Fig. 13.16 Coupled optical-power variation with the angular deviation with the number of iterations [78]

Figure 13.16 shows the curve of the coupling optical-power changing with the number of iterations under the influence of angular deviation. It can be observed that the coupling power of the single-mode optical fiber increases with an increase in the number of algorithm iterations, which means that the angular deviation is effectively compensated.

The SPGD algorithm can automatically align and position the fiber in a relatively short time, and the coupling efficiency reaches 53.2% [78]. Moreover, the coupling device has the advantage of a small and compact structure.

Mode-Conversion Method

Under the influence of atmospheric turbulence, some beam modes in optical wireless communication systems degenerate from a low-order mode to a high-order mode during transmission. This leads to a reduction in the fundamental mode energy, and thus, reduces the coupling efficiency of the single-mode fiber. Using a mode-conversion method, Zhang et al. overcame the above problems through mode-matching conversion and improved the efficiency of coupling spatial light to single-mode fibers [76, 84]. Figure 13.17 shows the optical path constructed using a liquid-crystal-spatial light modulator (LC-SLM) for mode conversion.

The mode-conversion results collected by the beam analyzer are shown in Fig. 13.18. In the figure, part of the mode of the second behavior spot degenerates into a higher-order mode after passing through atmospheric turbulence, and changes to LP01 mode after a mode conversion. It can be seen that, compared with the ideal light intensity distribution, the energy distribution of the LP02-mode light field in the first column is uniform and symmetrical; however, high-order diffraction components are still present.

The degraded higher-order mode is processed by LC-SLM2, and the obtained mode light field conforms to the mode field distribution under the ideal state. When the mode order is converted to a low order, the light field distribution is concentrated,

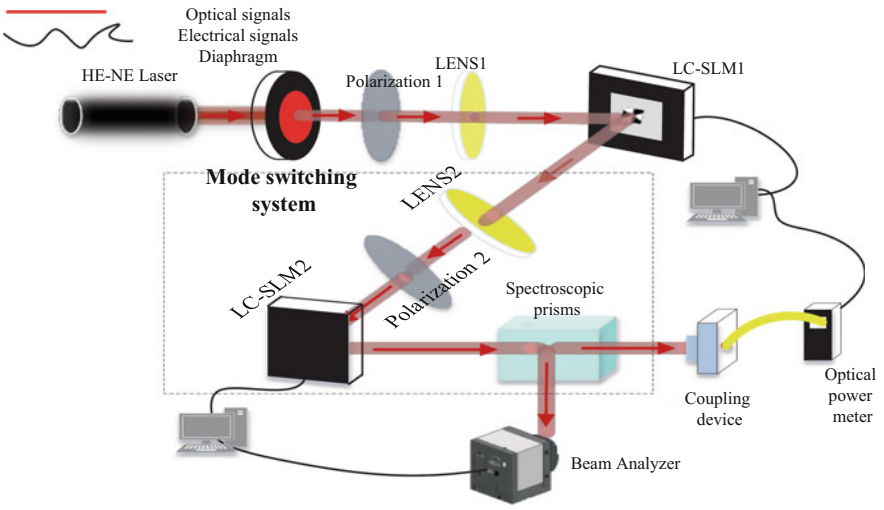


Fig. 13.17 Optical path diagram of a spatial-optical mode conversion based on LC-SLM [76]

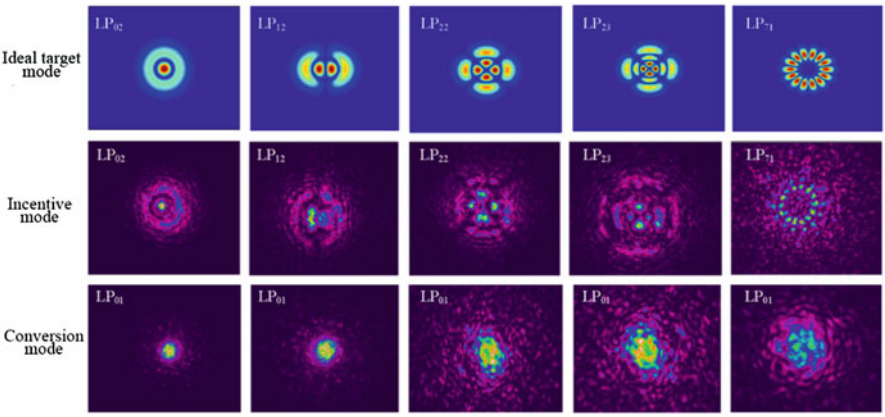


Fig. 13.18 Results of the schema transformation [85]

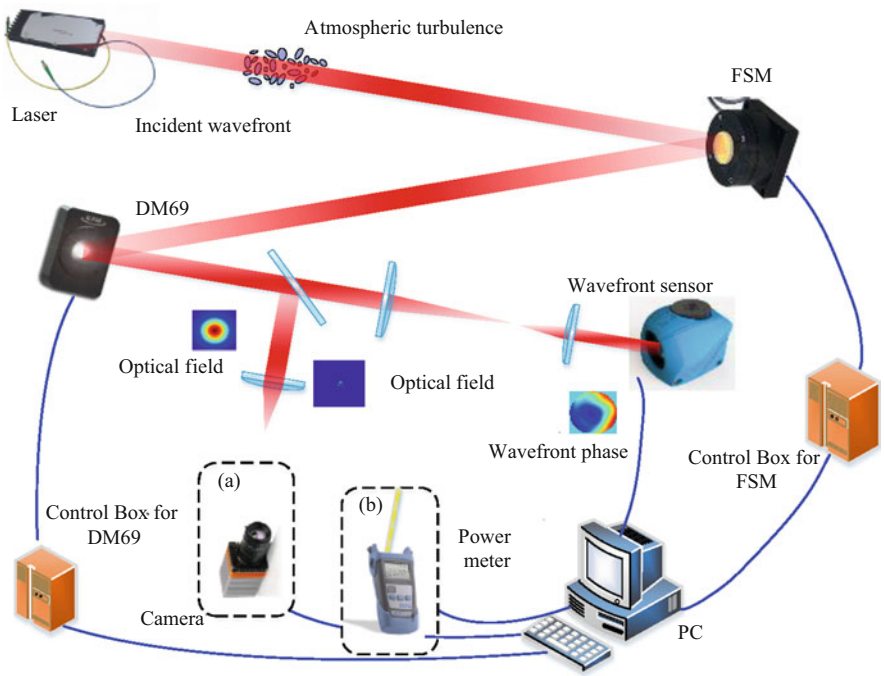


Fig. 13.19 Schematic diagram of an adaptive-optical wavefront correction fiber-coupled control system [77]

and the energy is located in the center of the light field, so the mode-conversion method plays an obvious role.

After pretreatment, the mode degenerates to LP22 when the coupling power is −18.53 dBm, and the coupling efficiency can be further increased from 55.9–66.2%

by using the simulated-annealing algorithm. It can be seen that the mode-conversion method plays an important role in improving the coupling efficiency. The coupling efficiency can be increased to 69.2% after converting the minimum low-order mode LP02. The coupling efficiency of the maximum high-order mode LP71 can be increased to 60.3%.

Adaptive-Optical Wavefront Correction

Because the efficiency of coupling a spatial light and fiber is reduced by the wavefront distortion caused by atmospheric turbulence, Yang et al. effectively improved the coupling efficiency of free-space coherent-optical communication by adopting adaptive-optics (AO) technology.

The adaptive-optical wavefront correction fiber-coupled control system is shown in Fig. 13.19. The adaptive-optical system consists of a wavefront sensor (WFS), deformable mirror (DM), bias fast-steering mirror (FSM), and personal computer (PC). The optical-fiber coupling system is composed of a coupling lens, beam splitter, optical-power meter, single-mode fiber, and infrared camera. The optical-fiber coupling power is controlled by the wavefront closed-loop control of the adaptive-optics system.

In an actual adaptive-optical system, owing to the static aberration of each lens, the wavefront phase located on the optical-fiber end face differs from that collected by the wavefront sensor. Therefore, the non-common optical-path aberration is usually corrected using the SPGD optimization algorithm without a wavefront mode. A closed loop consisting of a deformable mirror, optical-power meter, and computer was constructed.

Figure 13.20 shows the coupled optical-power curve and wave-front phase peak-to-valley (PV) curve before and after error elimination using the SPGD algorithm. It can be seen that before the error is eliminated, the power of the adaptive-optical closed loop decreases with an increase in the number of samples, while the coupling power increases with an increase in the number of samples after the error is eliminated, and finally stabilizes at −28.5 dBm.

Figure 13.21 shows the coupled power and PV curves after correcting the wavefront distortion. It can be seen that the coupled optical power increases from −35 to −27 dBm when the wavefront distortion is corrected, and the PV value of the wavefront is reduced from 5 to 1 μm. Figure 13.22 shows the coupling-power curves of the closed and open adaptive-optics loops. It can be seen from the figure that the coupling power increases from −33.75 to −30 dBm after correction, and the coupling capacity is significantly improved.

Figure 13.23 shows the coupling efficiency variation-trend curve with the increase of iteration times at various distances, according to the measured wavefront Zernike coefficient. As can be seen from the figure, after the wavefront correction, light waves are transmitted successively through distances of 500 m, 1 km, 5 km, 10 km, 100 km, and indoors.

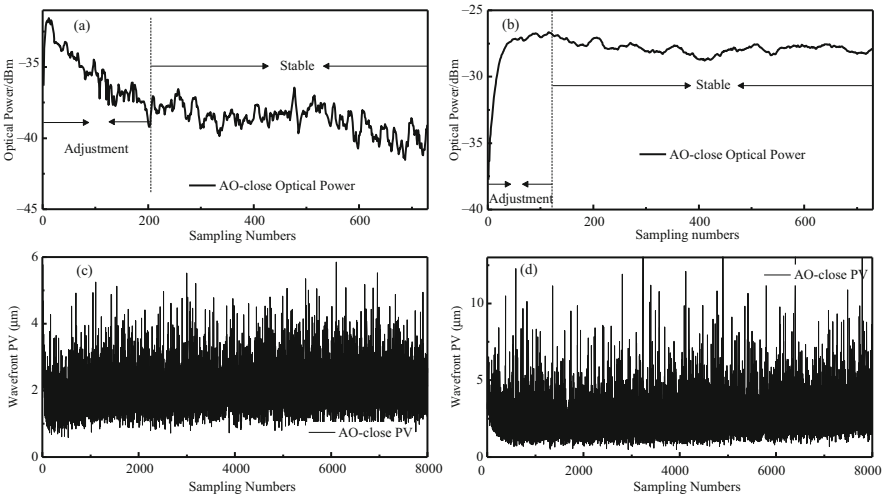


Fig. 13.20 Influence of non-common optical-path aberrations on AO closed-loop coupling power and wavefront phase [77]

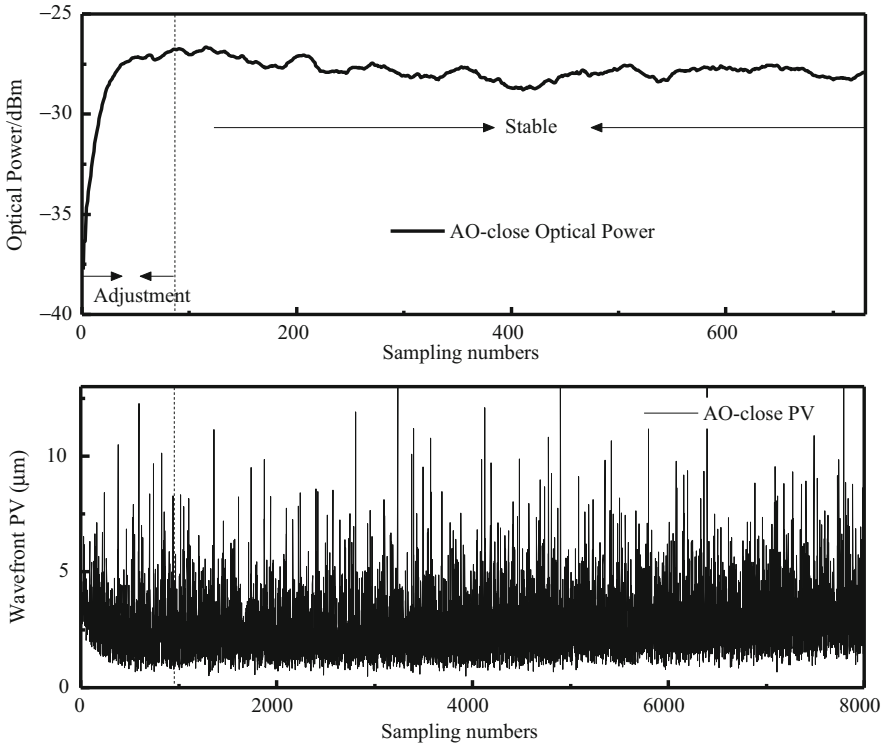


Fig. 13.21 Change curve and PV curve of fiber-coupling power before and after wavefront correction [77]

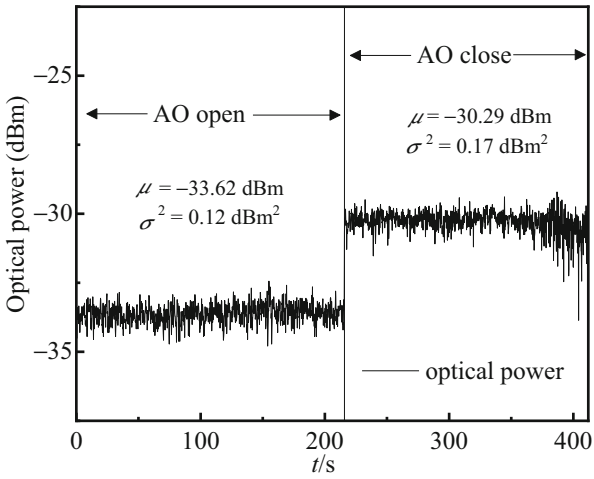


Fig. 13.22 Change curve of fiber-coupling power before and after wavefront correction [77]

In addition, it can be seen that with an increase in the communication distance, the fluctuation of the correction curve intensifies. The lifting effect of the long-distance correction is inferior to that of the short-distance correction, and the value fluctuation of the long-distance coupling efficiency is greater than that of the short-distance coupling efficiency. Therefore, under the influence of various factors, improving the efficiency of coupling long-distance spatial light to fiber is a complicated process, which is related to the beam alignment, beam mode, fiber-alignment error, wavefront measurement accuracy, detection speed, closed-loop correction bandwidth, and atmospheric-measurement environment.

Correction of Non-common Path Aberrations by Reverse-Transmission Calibration

Reverse-Transmission Calibration of a Non-common Optical-Path Aberration

Assuming that the beam-transmission direction is that shown in Fig. 13.19, and the phase introduced by atmospheric turbulence is $\varphi_{turbulence}$, the Zernike polynomial can be used to describe it in polar coordinates.

$$\varphi_{\text{turbulence}}(r,\theta) = \sum_{i=1}^{30} a_i \cdot z_i(r,\theta) \tag{13.1}$$

In Eq. (13.1), a_i is the Zernike coefficient and z_i is a Zernike series. *Noll* gives the relationship of the covariance $E(a_i, a_{i'})$ between Zernike coefficients a_i and $a_{i'}$ and the atmospheric coherence length r_0 [86].

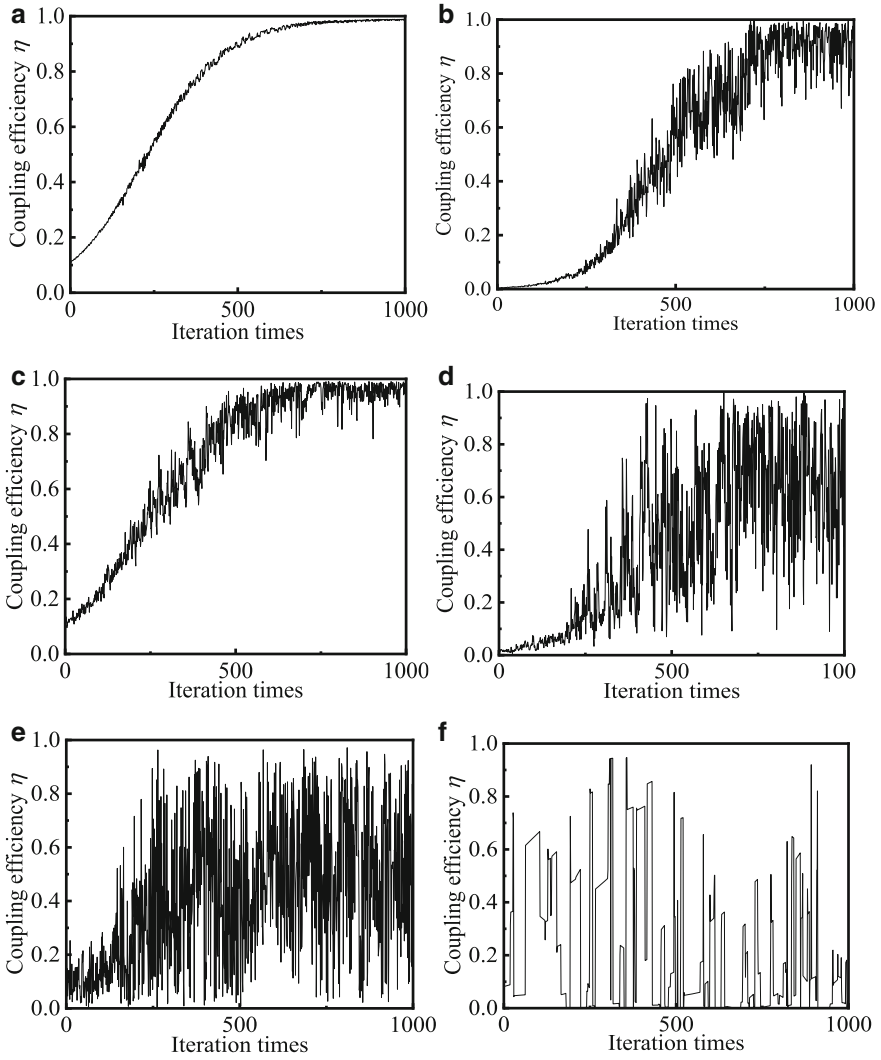


Fig. 13.23 Influence of the wavefront correction on the coupling efficiency at different distances. (**a**) indoor. (**b**) 500 m. (**c**) 1 km. (**d**) 5 km. (**e**) 10 km. (**f**) 100 km

$$E(a_i, a_{i'}) = \frac{2.2698(-1)^{(n+n'-2m)/2}\sqrt{(n+1)(n'+1)}\cdot\delta_Z\cdot\Gamma\left[\left(n+n'-\frac{5}{3}\right)/2\right]\cdot(D/r_0)^{5/3}}{\Gamma\left[\left(n-n'+\frac{17}{3}\right)/2\right]\cdot\Gamma\left[\left(n'-n+\frac{17}{3}\right)/2\right]\cdot\Gamma\left[\left(n+n'+\frac{23}{3}\right)/2\right]} \tag{13.2}$$

In Eq. (13.2), n and n' are the radial order of coefficients a_i and $a_{i'}$, respectively; m is the angular order of coefficient a_i; z is the Kronecker function, and D is the

aperture of the receiving antenna. The aberrations existing in lenses 3, 4, and 5 are φ_3, φ_4, and φ_5, respectively, and the aberrations generated by the transmission and reflection of the spectroscope are φ_T and φ_R. The phase information φ_{mea} collected by the branch of the wavefront sensor can be expressed as

$$\varphi_{\text{mea}} = \varphi_{\text{IWF}} + \varphi_{\text{turbulence}} + v \cdot IF_{\text{cmd2phase}} + \varphi_{\text{DM-error}} + \varphi_T + \varphi_3 + \varphi_4. \quad (13.3)$$

In Eq. (13.3), φ_{IWF} is the phase of the outgoing laser wavefront after collimation, v is the voltage signal applied by the deformable mirror, $IF_{\text{cmd2phase}}$ is the influence function of the surface shape of the deformable mirror, and $\varphi_{\text{DM-error}}$ is the plane-fitting error of the deformable mirror, which can be fitted by a Gaussian function.

$$\varphi_{\text{DM-error}}(x, y) = \exp\left(\ln(\omega)\left(\frac{\sqrt{(x - x_i)^2 + (y - y_i)^2}}{d}\right)^{\alpha}\right), i = 1, 2, \cdots n \quad (13.4)$$

In Eq. (13.4), α stands for the Gaussian index, d is the distance between the individual drives, (x_i, y_i) is the coordinate value of the ith drive, n is the total number of drives, and ω represents a cross-link value. At this time, the wavefront aberration generated by the φ_{focal} plane of the communication branch coupled lens (i.e., the coupling fiber end face) can be expressed as

$$\varphi_{\text{focal}} = \varphi_{\text{IWF}} + \varphi_{\text{turbulence}} + v \cdot IF_{\text{cmd2phase}} + \varphi_{\text{DM-error}} + \varphi_R + \varphi_5. \quad (13.5)$$

Equations (13.5) and (13.3), and the difference of $(\varphi_T + \varphi_3 + \varphi_4) - (\varphi_R + \varphi_5)$ are known as the common optical-path aberration.

To measure and correct non-common optical-path aberrations, the traditional method is to adopt an optimization algorithm represented by the SPGD algorithm. The loop in Fig. 13.19, consisting of a deformable mirror, optical-power meter, and computer, is taken as a wavefront-free sensing system and passed through the iteration meter [87].

$$v(k + 1) = v(k) + \gamma \cdot \Delta v(k) \cdot \Delta P(k) \quad (13.6)$$

In Eq. (13.6), $v(k)$ is the voltage applied to the deformed mirror in the kth iteration, ΔP is the change in the objective-function value of the system (coupled optical power) in the kth iteration, γ is the disturbance amplitude, and $\Delta v(k)$ is the control quantity of the random disturbance applied to each driver of the deformed mirror in the kth iteration, which follows both the independent and Bernoulli distributions. When the coupling optical power reaches convergence stability, the wavefront phase acquired by the wavefront sensor at this moment is denoted as $\varphi_{\text{mea-S}}$.

Reverse transmission means that an identical beam is sent from the receiving end to the sending end. We assume multiple emission light sources into the laser-coupling optical-fiber access, as shown in Fig. 13.24. The position of the deformable

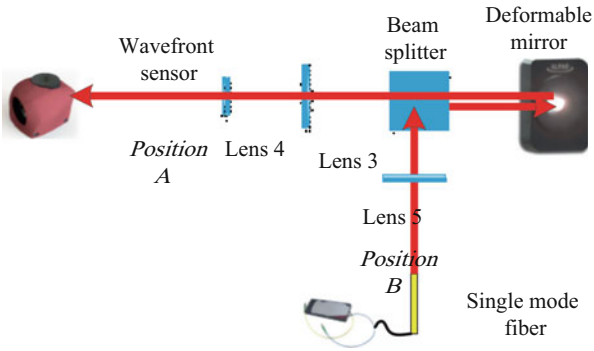


Fig. 13.24 Measurement of a non-common path aberration using the reverse-transmission calibration method

mirror is 10° counterclockwise. The deformable mirror is instructed to simultaneously clear the conditions and the reflection is in a fully planar deformation state. After the light output by the coupled fiber is reflected by the spectroscope, it is reflected by the deformable mirror directly into the wavefront sensor, to measure the wavefront information at the current moment.

As shown in Fig. 13.24, in the reverse-transmission state, the wavefront phase output by the fiber-coupling port is denoted as $\varphi_{\text{fiber-I}}$. According to the optical-path reversibility principle, the optical field carrying the phase information $\varphi_{\text{fiber-I}}$ can produce the highest coupling efficiency. At this time, the wavefront phase collected by the wavefront sensor can be expressed as

$$\varphi_{\text{mea-I}} = \varphi_{\text{fiber-I}} - \varphi_5 - \varphi_R + \varphi_{\text{DM-error}} + \varphi_T + \varphi_3 + \varphi_4. \tag{13.7}$$

Because $\varphi_{\text{mea-I}}$ contains both, it maximizes the coupling efficiency of the wavefront phase $\varphi_{\text{fiber-I}}$ and also includes the common optical path as the differential phase $(\varphi_T + \varphi_3 + \varphi_4) - (\varphi_R + \varphi_5)$.

Adaptive-Optical System

A block diagram of the adaptive-optical closed-loop control with wavefront sensing shown in Fig. 13.19 is shown in Fig. 13.25. The nth closed-loop iteration can be expressed as

$$v(k+1) = v(k) + k_i \cdot [\varphi_{\text{tar}} - \varphi_{\text{mea}}(k)] \cdot CM_{\text{phase2cmd}}, \tag{13.8}$$

where φ_{tar} controls the target wavefront phase, φ_{mea} is the phase collected by the wavefront sensor, $CM_{\text{phase2cmd}}$ is the command matrix for transforming the wavefront phase to the voltage of the deformable mirror, and k_i represents the integral gain. Considering the algorithm complexity and calculation speed of a real-time control system, the wavefront can be controlled by a single integral operation.

According to the mode field matching principle and Fresnel diffraction formula, the discretization expression of the fiber-coupling efficiency can be expressed as

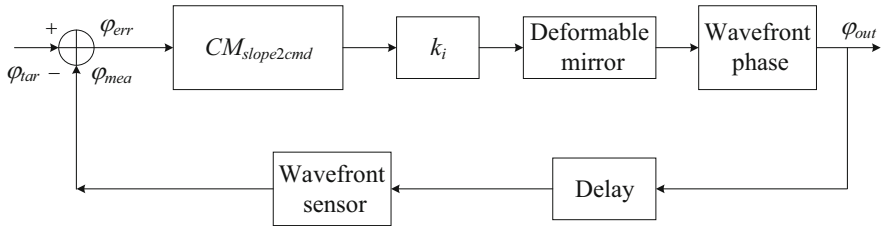


Fig. 13.25 Block diagram of an adaptive-optical closed-loop control

$$\eta = \frac{\left|\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}\frac{W_m}{M}\cdot\frac{2\pi}{N}\cdot U_f(u,v)\cdot U_i^*(u,v)\cdot\frac{u}{M}W_m\right|^2}{\left(\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}\frac{W_m}{M}\cdot\frac{2\pi}{N}\cdot U_f(u,v)\cdot U_f^*(u,v)\cdot\frac{u}{M}W_m\right)\left(\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}\frac{W_m}{M}\cdot\frac{2\pi}{N}\cdot U_i(u,v)\cdot U_i^*(u,v)\cdot\frac{u}{M}W_m\right)}, \tag{13.9}$$

where M and N are the discrete points of the radial and angular coordinates, respectively; and W_m is the mode field radius of the single-mode fiber. U_f is the mode field distribution of the single-mode fiber [88]:

$$U_f(u,v) = \sqrt{\frac{2}{\pi}}\frac{1}{W_m}\cdot\exp\left[-\left(\frac{u}{M}\right)^2\right]. \tag{13.10}$$

U_i is the optical-field distribution located at focal plane 5 of the lens (i.e., the end face of the coupled fiber), which can be expressed as [88]

$$U_i(u,v) = \frac{\exp(ikf)}{ikf}\cdot\exp\left(\frac{ik\left(\frac{u}{M}W_m\right)^2}{2f}\right)\cdot\left(\frac{D_a}{2}\right)^2\cdot\phi(u,v), \tag{13.11}$$

where f_5 is the focal length of lens 5, D_a is the diameter of lens 5, $k = 2\pi/\lambda$ is the wave vector, and (u, v) is the discrete two-dimensional Fourier transform of φ_{focal} [93]:

$$\phi(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\varphi_{\text{focal}}(x,y)\cdot\frac{2}{MN}\cdot\frac{x}{M}\cdot\exp\left[-j2\pi\left(\frac{x}{M}W_m\right)\cdot\left(\frac{u}{M}\frac{D_a}{2\lambda f}\right)\cdot\cos\left(\frac{y}{N}2\pi-\frac{v}{N}2\pi\right)\right].$$
$$u = 0,1,\cdots,M-1$$
$$v = 0,1,\cdots,N-1 \tag{13.12}$$

Simulation Analysis

Numerical simulations were performed according to Eqs. (13.1–13.12). The integral coefficient $k_i = 0.3$, $\varphi_{\text{IWF}} = 0$, M and N were both 50, $W_m = 5.25$ m, $D_a = 25.4$ mm,

$\lambda = 1550$ nm, and $CM_{\text{slope2cmd}}$ and $IF_{\text{cmd2phase}}$ adopted the surface matrix provided by ALPAO, an adaptive-optics product manufacturer. The Gaussian index of the deformable mirror was $\alpha = 2$, the cross-linking value was $\omega = 0.08$, and the variance was 10^{-4}. The drive spacing $d = 1.5$ mm. Lenses 3, 4, and 5 are flat convex lenses with focal lengths of 175, 75, and 125 mm, respectively.

According to the integral-conversion relationship between geometric and wave aberrations, the differences in the wave phase after the beam passes through the optical device are calculated as φ_3, φ_4, and φ_5. The aberrations of the transmitted and reflected waves of the spectroscope are denoted as φ_T and φ_R, respectively. $\varphi_{\text{turbulence}}$ follows the Karhunen–Loeve function. $D/r_0 = 1$ is used to calculate the phase of the Zernike coefficient.

When $\varphi_{\text{tar}} = 0$, that is, the phase of the wavefront collected by the wavefront sensor is controlled as a plane wave, the peak–valley (PV) value and root-mean-square (RMS) value of the wavefront in the uncorrected and closed-loop states, as well as the coupling efficiency, can be obtained according to Eq. (13.8).

With a non-common optical-path aberration, after 1000 frames of closed-loop control iteration, the PV values of the wavefront decrease from an uncorrected 18.59 m to a closed loop 1.78 m, and the RMS value of the wavefront decreases from an uncorrected 1.51 m to a closed loop 0.24 m. The three-dimensional figures in Fig. 13.26a represent the wavefront phase collected by the wavefront sensor in the initial state (left) and after the closed-loop control of 1000 frames (right), according to the arrow direction. The wavefront flatness after the closed loop was obviously better than that in the uncorrected state. As shown in Fig. 13.26b, the coupling efficiency decreased from 36.52% to 18.79%. This is because $\varphi_T + \varphi_3 + \varphi_4$ in Eq. (13.3) is not equal to $\varphi_R + \varphi_5$ in Eq. (13.5).

$\varphi_{\text{tar}} = \varphi_{\text{mea-S}}$ and $\varphi_{\text{tar}} = \varphi_{\text{mea-I}}$. The SPGD algorithm and reverse-transmission calibration method were adopted to measure and correct the non-common optical-path aberrations. The disturbance amplitude $\gamma = 0.6$, the random disturbance control quantity $\Delta v = 0.01$ from the Bernoulli distribution, and φ_{tar} were used to reconstruct

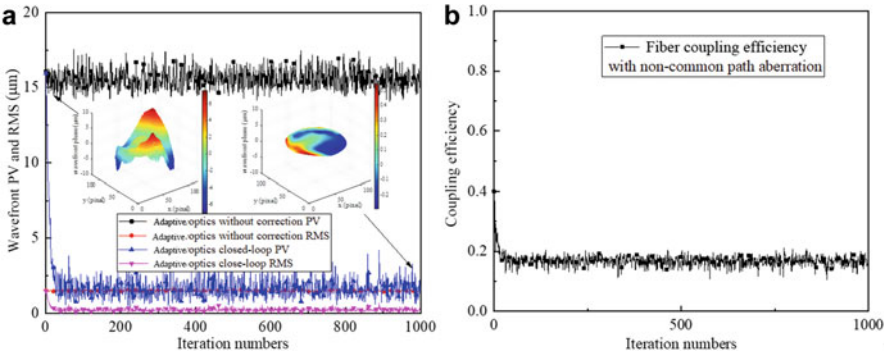


Fig. 13.26 Adaptive-optical system with non-common path aberrations. (**a**) Uncorrected and closed-loop PV values and RMS values of the wavefront. (**b**) Closed-loop coupling efficiency

the wavefront. The PV and RMS values of the wavefront and coupling efficiency in the uncorrected and closed-loop states were obtained, as shown in Fig. 13.28.

The SPGD algorithm was used to measure and correct the non-common optical-path aberration. After the closed loop, the PV value of the wavefront decreased from 18.55 to 1.75 μm, the RMS value of the wavefront decreased from 1.51 to 0.24 μm, and the coupling efficiency increased from 36.79–48.41%. This is because the SPGD algorithm is affected by turbulence, as shown in Fig. 13.27a, so it easily falls into local extrema; thus, it fails to fully calculate the wavefront phase information of the non-common optical-path aberration when the iteration is stopped. The maximum objective function determines the coupling efficiency in the closed-loop state. The PV value of the wavefront decreased from 18.64 to 3.28 μm, and the RMS value decreased from 2.76 to 0.68 μm.

As shown in the inset figures in Fig. 13.24a, c, there is no significant difference between the uncorrected and corrected wavefront phases, and the convergence degree is the same. The corresponding conversion was performed in the reconstruction process after closed-loop control, and the coupling efficiency of the reverse-

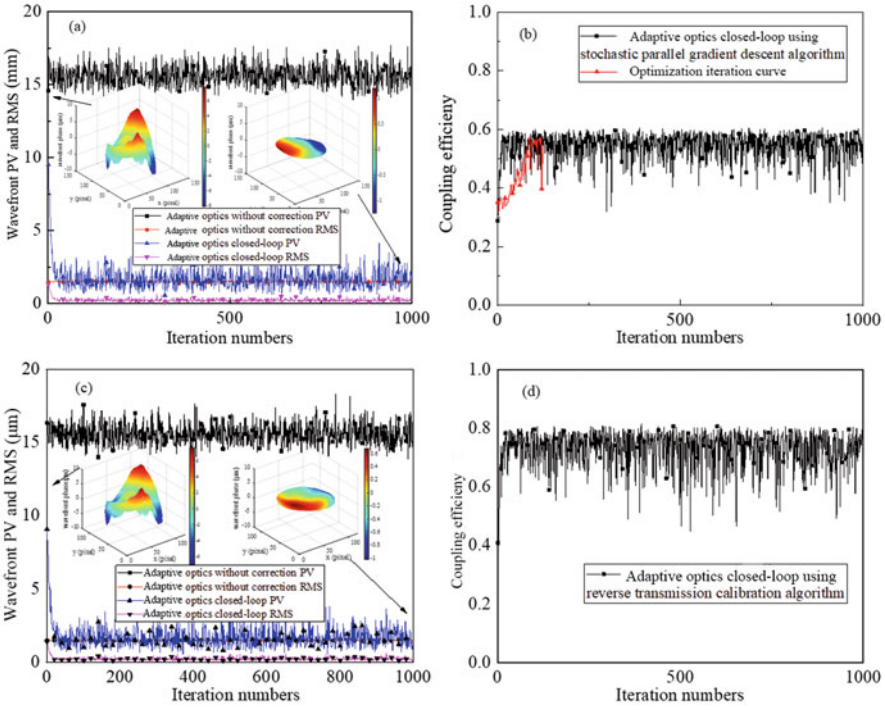


Fig. 13.27 Adaptive-optical system for correcting non-common path aberrations. (**a**) Uncorrected and wavefront RMS and PV values of the SPGD method. (**b**) Coupling efficiency of SPGD method. (**c**) Coupling efficiency of uncorrected and wavefront RMS and PV values of the reverse-transmission calibration method. (**d**) Coupling efficiency of reverse-transmission calibration method

transmission calibration method improved from 28.31–69.14%. This is because the beam-transmission direction in Fig. 13.27c is not affected by the turbulent environment, and the non-common optical-path aberration can be measured completely and steadily. In the closed-loop state, the coupling efficiency is better improved than that of the SPGD algorithm.

Experimental Results and Analysis

(a) **Measurement and correction of non-common path aberrations in a static environment**

The local oscillator light in a optical wireless coherent communication system was connected to a single-mode fiber using the reverse-transmission calibration method. Figure 13.28 shows the wavefront slope and corresponding wavefront phase of the non-common optical-path aberration measured by the wavefront sensor. The PV value is 3.71 μm and the RMS value of the wavefront is 1.34 μm, which are sufficient to significantly affect the coupling efficiency.

Figure 13.29 shows the coupling efficiency of the adaptive-optical system without correcting the non-common optical-path aberration, using the SPGD and reverse-transmission calibration methods, and the PV and RMS values of the wavefront before and after conversion.

The sampling rate of the wavefront sensor after reconstructing the wavefront (approximately 50 Hz) was approximately ten times than that of the serial-port sampling rate of the optical-fiber power meter (approximately 5 Hz). The coupling efficiency decreased from 12.83% to 6.79% after the closed loop without correction. Using the SPGD method, the coupling efficiency increased from the initial 12.88–25.28%. The coupling efficiency of the reverse-transmission calibration method increased from 10.04–45.21%.

The PV values of the wavefront without conversion reconstruction tend to first decrease and then increase. This is because the wavefront phase of the fiber branch

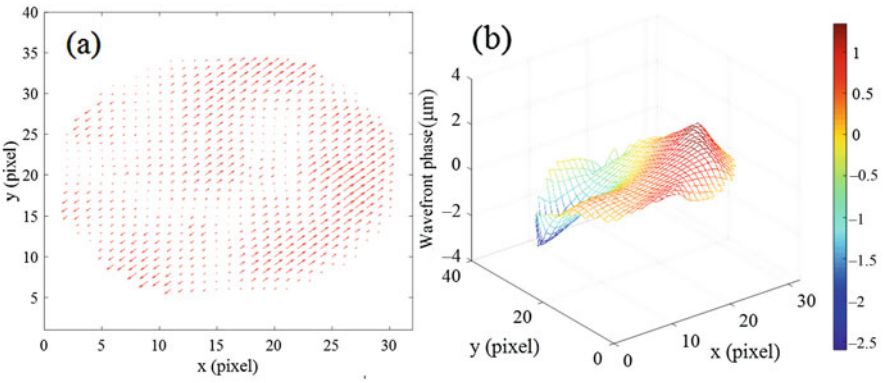


Fig. 13.28 Non-common optical-path aberration. (**a**) Wavefront slope; (**b**) wavefront phase in the actual system

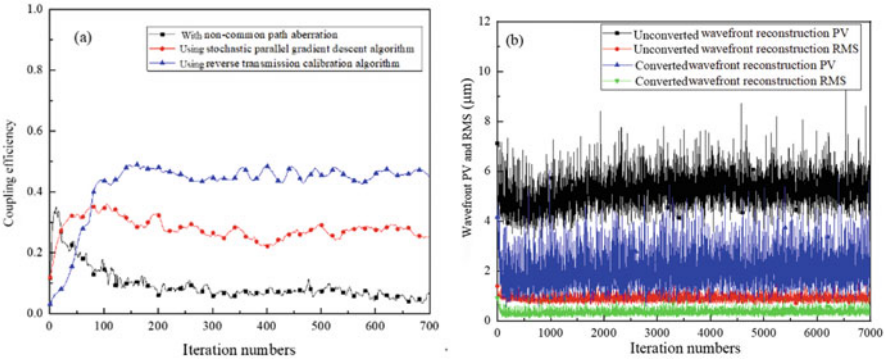


Fig. 13.29 1550-nm band adaptive-optical system closed-loop measurement. (**a**) Coupling efficiency of the uncorrected non-common optical-path aberration and the SPGD and reverse-transmission calibration methods; (**b**) unconverted and converted wavefront PV values and RMS values

reaches the optimal coupling efficiency after correction from the initial state. The wavefront-sensing branch corresponding to the same time period does not show the same numerical trend, owing to the aberration conversion of the non-common optical path and the positive and negative characteristics of the wavefront phase value. In addition, the internal noise of the wavefront sensor leads to some synaptic data in the wavefront slope measurement and wavefront phase calculation, but does not significantly affect the fluctuation of the coupling efficiency.

(b) **Field experiment in a turbulent environment**

An outfield experiment for the coherent detection of optical wireless coherent communication was conducted. The two ends of the communication were located at the South Second Ring Road of Xi'an City and the sixth floor of Xi'an University of Technology, respectively. The communication distance was 1.3 km, and the measurement date was April 16, 2021; the weather was clear with a level-2 southwest wind.

The structure of the optical wireless coherent communication system using adaptive optics is shown in Fig. 13.30. The transmitter uses external modulation to load the signal source onto the modulator. After amplification by the fiber amplifier, the optical antenna collimates the output. The output power of the fiber amplifier is 16.77 dBm.

After the beam is converged by the receiving antenna, adaptive-optics technology is used for wavefront correction, and the spatial light is directly coupled into the single-mode fiber. After frequency mixing and balance-detection processing, the intermediate-frequency (IF) signal is output and demodulated. The mixer model is Kylia COH24, the balancing detector model is General Photonics BPD-002, and the signal laser and local oscillator are both Koheras BASIK E15 narrow-width fiber lasers.

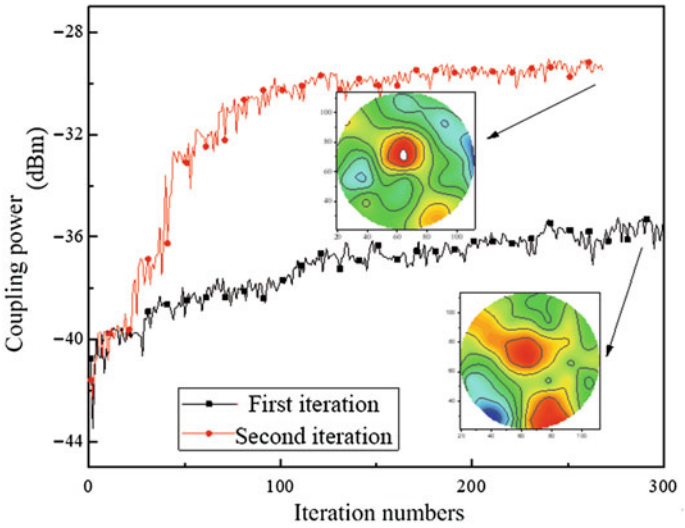


Fig. 13.30 Coupling-power variation and corresponding deformed-mirror profiles under different iterations using the SPGD method

The coupled light power was taken as the evaluation index, and the SPGD algorithm was adopted to obtain the coupled power at different times. The deformation mirror shape was fitted according to the voltage applied by the deformation mirror when the iteration was stopped, as shown in Fig. 13.30.

As shown in Fig. 13.30, the maximum power values obtained by the different iterations differ. The inset figures in Fig. 13.30 show obvious defocusing terms in the surface shapes fitted by the deforming mirror when the iteration is stopped; however, some differences still exist. The accuracy of the aberration measurement in a non-common optical path is bound to be affected by the local optimization of the different SPGD profile errors.

Figure 13.31 shows the non-common optical-path aberrations measured by the SPGD algorithm at different times. The wavefront phase presents a ring structure because the secondary mirror of the receiving antenna blocks the received light spots, and the measurement values of the non-common optical-path aberration in Fig. 13.31a are different from those in Fig. 13.31b. The measurement results at different times are also affected by atmospheric turbulence and thus differ.

Figure 13.32 shows the coupling efficiency curve and IF-signal waveform in the uncorrected and closed-loop adaptive-optics states after measuring and correcting non-common optical-path aberrations using the reverse-transmission calibration method. The adaptive-optics coupling efficiency in the uncorrected and closed-loop states increased from 19.72% to 36.93%. The peaking value of the IF signal increased from 400 mV to 1.4 V, while the maximum output peaking value of BPD-002 is 3.1 V. The mixing efficiency of the system increased from 1.66% in the uncorrected state to 20.40% in the closed-loop state.

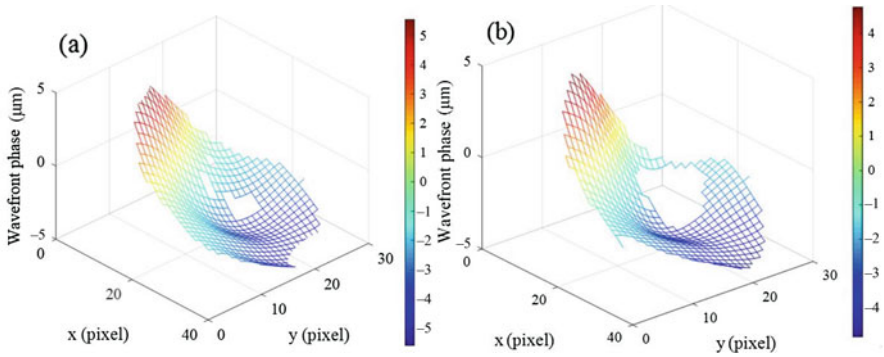


Fig. 13.31 1.3-km experimental link using the SPGD algorithm to measure a non-common optical-path aberration. (**a**) First measurement; (**b**) second measurement

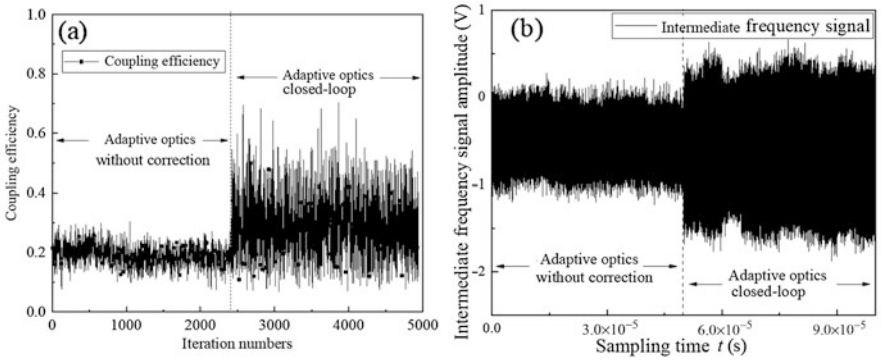


Fig. 13.32 Uncorrected and closed-loop states of 1.3-km adaptive optics in a field experiment. (**a**) Coupling efficiency; (**b**) IF signal waveform

(c) **Field experiment in a complex environment**

Based on the above experiments, the transmitting end is located in Xiaozhai Village, Bailuyuan, Xi'an City, and the receiving end is located on the sixth floor of Xi'an University of Technology. At the receiving end, a piezoelectric deformable mirror is used to randomly generate wavefront distortions of different frequencies and amplitudes; then, the wavefront is corrected and the fiber is coupled through the adaptive-optical system. The experiment took place on November 6, 2021, when the weather was cloudy with light rain. The experiment-link distance was 10.2 km.

The transmitter output power of the fiber amplifier is 23.01 dBm. Before correction, the power of the receiver coupled into the single-mode fiber fluctuated from −44 to −27 dBm. The coupling power does not increase significantly with the transmission-power increase; thus, it is necessary to use adaptive-optics technology to improve and stabilize the coupling efficiency.

Figure 13.33 shows the coupled-power iteration curve obtained using the SPGD algorithm for the 10.2-km link and the wavefront phase collected by the wavefront

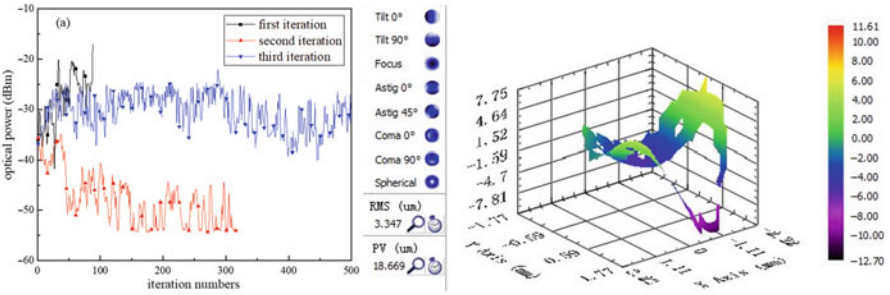


Fig. 13.33 10.2-km link adaptive-optical initialization using the SPGD algorithm. (**a**) Coupled optical-power iteration curve; (**b**) wavefront phase

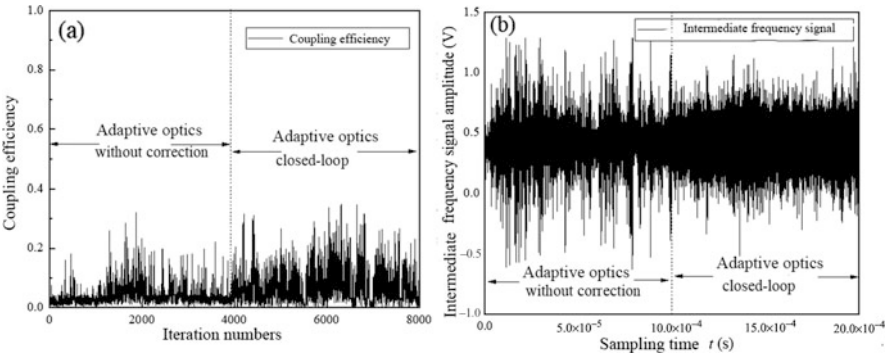


Fig. 13.34 Uncorrected and closed-loop states in a 10.2-km adaptive-optics field experiment. (**a**) Coupling efficiency; (**b**) IF signal waveform

sensor when the iteration stopped. The coupling power does not obviously increase with the increase of the number of iterations, and some of it even declines. The wavefront phase is difficult to detect, owing to the light intensity flicker caused by atmospheric turbulence. Thus, the non-common path aberration-information acquisition is incomplete. This indicates that the SPGD algorithm is not universal for adaptive-optical initialization calibration over long-distance transmissions.

Figure 13.34 shows the coupling efficiency and IF-signal waveform in the closed-loop state after adaptive-optics initialization with uncorrected adaptive optics and the reverse-transmission calibration method on a 10.2-km link. The average coupling efficiency increased from 3.91% in the uncorrected state to 10.13% in the closed-loop state, while the variance of the peak-to-peak value of the IF signal increased from 33 mV in the uncorrected state to 123 mV in the closed-loop state. The average mixing efficiency of the system increased from 3.87% in the uncorrected state to 8.52% in the closed-loop state. This indicates that with the increase in communication distance, atmospheric turbulence has a more significant effect on the adaptive-optical correction effect than the non-common optical-path aberration.

According to the reversible optical-path principle, a calibration method using a reverse transmission is proposed to measure and correct the aberration of a non-common optical path. The closed-loop control of the non-common optical-path aberration improves the coupling efficiency while correcting the wavefront. Compared with the traditional SPGD algorithm, this scheme does not fall into local optima, is not affected by the external environment, and can assist in the alignment and debugging of the optical path. It is simple and feasible, easy to implement in engineering, and has reference significance and practical value for the fiber-coupling technology of optical wireless coherent communication systems.

Comparison of Several Different Methods to Improve the Spatial Optical-Fiber Coupling Efficiency

Comparison of Different Methods to Improve Fiber Coupling

Table 13.5 compares the research on three methods, namely automatic alignment, mode conversion, and adaptive-optical wavefront correction. They improve the

Table 13.5 Comparison of different methods to improve the spatial optical-fiber coupling efficiency

Method classification		Applicable background	Method	Coupling efficiency
Automatic-alignment method	Simulated-annealing algorithm	The single-mode optical-fiber core is small, the manual alignment is time consuming and laborious, and the precision is low	The voltage on the two-dimensional piezoelectric ceramic is controlled to drive the optical-fiber displacement, and the optical-fiber position is adjusted in real time	Reached 51.4%
	SPGD algorithm		The optical-fiber position is adjusted in real time by controlling the voltage on the 5-DOF piezoelectric ceramic	Reached 53.2%
Mode-conversion method		Owing to the influence of atmospheric turbulence, the beam mode degrades and the energy of the fundamental mode decreases	The coupling efficiency of the spatial-optical single-mode fiber is improved by mode conversion	Maximum was 69.2%
Adaptive-optical wavefront correction method		The wavefront distortion caused by atmospheric turbulence reduces the efficiency of coupling the space light and fiber	Adaptive-optics technology is used to effectively improve the coupling efficiency of the space-optical single-mode fiber	Maximum increase was 80.0% at different distances

coupling efficiency of spatial optical fiber from different aspects, such as method classification, application background, specific methods, and coupling efficiency.

Comparison of Domestic and Foreign Research on the Automatic-Alignment Method

Table 13.6 compares the automatic-alignment research results of Xi'an University of Technology to those of Adept Company [89] and Zhejiang University [90], in terms of the alignment platform, control method, automatic-alignment algorithm, and coupling efficiency.

Comparison of Domestic and Foreign Research on Adaptive-Optical Wavefront Correction

Table 13.7 compares the research results of the adaptive-optical wavefront correction method of Xi'an University of Technology to those of Lawrence Livermore Laboratory, Jilin University and Changchun University of Science and Technology, in terms of the correction algorithm, transmission distance, and coupling efficiency. It provides some theoretical basis and reference value for the future research and development in this field.

Table 13.6 Comparison of domestic and foreign automatic-alignment methods

Researcher	Alignment platform	Control method	Automatic-alignment algorithm
Adept Company (U.S.)	Three degrees of freedom x, y, z	Optical-power feedback	Rough alignment: Image processing Fine alignment: Mountain-climbing method
Zhejiang University	3D micro-workbench: x, y, z	Optical-power feedback	Sweep sampling method
Xi'an University of Technology	Five degrees of freedom x, y, z, θx, θy	Optical-power feedback	SPGD algorithm

Table 13.7 Comparison of domestic and foreign achievements in adaptive-optical wavefront correction methods

Researcher	Correction algorithm	Transmission distance (km)	Coupling efficiency
Lawrence Livermore Laboratory, USA	SPGD algorithm	28.2	33.0% increase
Jilin University	SPGD algorithm, no wavefront detection control algorithm	20.0	69.0% increase
Xi'an University of Technology	SPGD algorithm	11.0	35.2% increase
Changchun University of Science and Technology	Zernike mode-correction method	1.0	20.0% increase

Summary

This chapter systematically summarized the domestic and foreign research in the field of space optical-fiber coupling, and introduced the work in related fields by Xi'an University of Technology.

The first part of the chapter discussed overcoming the influence of atmospheric turbulence on the space optical-fiber coupling efficiency, such as improving the beam quality from the source, and optimizing and improving the coupling-system scheme. In the second part, the automatic-alignment, mode-conversion, and adaptive-optical wavefront correction methods were introduced to improve the coupling efficiency.

In simulations and experimental measurements, different technologies improved the coupling efficiency of spatial optical fiber to different degrees. This lays a foundation for development in this field, but also has many shortcomings.

1. Automatic alignment: The algorithm operating bandwidth is low and the piezoelectric-ceramic modules influence each other.
2. Mode conversion: The conversion can only be carried out between single modes, not multiple modes.
3. Adaptive-optical wavefront correction: The correction effect is not obvious for long distances.

In the future, it is necessary to develop more intelligent, convenient, efficient, and simpler spatial optical-coupling technologies, to send more effective optical-signal transmissions and improve the communication quality. To promote the development of spatial light-fiber coupling technology in optical wireless communication, the following areas have been investigated:

1. In the mode-conversion method, an excellent feedback optimization algorithm was studied to realize free-space light mode conversion and improve the fiber-coupling efficiency [75].
2. In the mode-conversion method, based on single-mode conversion, the simultaneous conversion of multiple modes was studied to achieve a better coupling effect [77].
3. In the automatic-alignment method, a more stable and compact optical-fiber structure based on piezoelectric ceramics was developed, which further improved the stability and accuracy of the optical-fiber position [78].
4. To optimize the coupling structure, a more convenient and accurate tuning structure was developed to realize an integrated-transceiver Maka antenna with higher coupling efficiency [82].
5. Automatic spatial light-fiber coupling alignment technology was investigated under different turbulence intensities and transmission models to improve the fiber-coupling efficiency under atmospheric turbulence [83].

References

1. Ke Xizheng, Optical wireless communication [M]. Beijing: Science Press, 2016.
2. Luo Jing. Research on spatial optical coupling auto-alignment technology [D]. Xi'an: Xi'an University of Technology, 2019.
3. Shaklan S, Roddier F. Coupling starlight into single-mode fiber optics [J]. Applied Optics, 1988, 27(11): 2083–2362.
4. Salisbury M S. Sensitivity and signal to noise ratio improvement of a one micron ladar system incorporating a neodymium doped optical fiber preamplifier. Laser radar testbed [J]. Nasa Sti/recon Technical Report N, 1994, 94(1): 1–3.
5. Winzer P J, Leeb W R. Fiber coupling efficiency for random light and its applications to lidar [J]. Optics Letters, 1998, 23(13): 986–989.
6. Ruilier C A. study of degraded light coupling into single-mode fibers [J]. Proceedings of SPIE The International Society for Optical Engineering, 1998, 3350(1): 319–3210.
7. Weyrauch T, Vorontsov M A, Gowens J W, et al. Fiber coupling with adaptive optics for free-space optical communication [C]//Free-Space Laser Communication and Laser Imaging. San Diego, California, United States, 22 January, 2002: 177–184.
8. Dikmelik Y, Davidson F M. Fiber-coupling efficiency for free-space optical communication through atmospheric turbulence [J]. Applied Optics, 2005, 44(23): 4946–4952.
9. Toyoshima M. Maximum fiber coupling efficiency and optimum beam size in the presence of random angular jitter for free-space laser systems and their applications [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2006, 23(9): 2246–2250.
10. Carriz C E, Calvo R M, Belmonte A. Proof of concept for adaptive sequential optimization of free-space communication receivers [J]. Applied optics, 2019, 58(20): 5397–5403.
11. Franz F, Oswald W. Application of Single-Mode Fiber-Coupled Receivers in Optical Satellite to High-Altitude Platform Communications [J]. Eurasip Journal on Wireless Communications & Networking, 2008, 2008(1): 1–8.
12. Takenaka H, Toyoshima M. Study on the fiber coupling efficiency for ground-to-satellite laser communication links [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2010, 7587(6): 1–13.
13. Takenaka H, Toyoshima M, Takayama Y. Experimental verification of fiber-coupling efficiency for satellite-to-ground atmospheric laser downlinks [J]. Optical Express, 2012, 20(14): 15301–15309.
14. Edwards B L, Israel D J, Vithlani S K. Latest Changes to NASA's Laser Communications Relay Demonstration Project [C]//SPIE Photonics West Conference. San Francisco, California, United States, 15 February, 2018: 204–214.
15. Edwards B L, Isreal D J, Caroglanian A, et al. A Day in the Life of the Laser Communications Relay Demonstration Project [C]//14th International Conference on Space Operations. Daejeon, Korea, 16–20 May, 2016: 1–16.
16. Boroson D M. Overview of Lincoln Laboratory development of lasercom technologies for space [J]. SPIE, 1993, 18(6): 30–310.
17. Sayano K, Nguyen I A, Chan J K. Demonstration of multichannel optical CDMA for free-space communications [C]//Free-Space Laser Communication Technologies XIII. San Jose, California, United States 20 June, 2001: 38–410.
18. Modavis R A, Webb T W. Anamorphic microlens for laser diode to single-mode fiber coupling [J]. IEEE Photonics Technology Letters, 1995, 7(7): 798–800.
19. Wallner O, Winzer P J, Leeb W R. Alignment tolerances for plane-wave to single-mode fiber coupling and their mitigation by use of pigtailed collimators [J]. Applied Optics, 2002, 41(4): 637–643.
20. Sherman P, John A T. Point-to-Multipoint Free-Space Optical Wireless Communication System [P]. US Patents, 6445496B1, 2002-9-3.
21. Daniel V H, David M B, Nathan W R, et al. Fiber optical bundle array wide field-of-view optical receiver for free space optical communications [J]. Optical letter, 2010, 35(21): 3359–3361.

22. Hottinger P, Harris R J, Dietrich P I, et al. Micro-lens array as tip-tilt sensor for single-mode fiber coupling [C]//Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation III. Austin, Texas, United States, 10 July, 2018: 617–631.
23. Du K, Baumann M, Ehlers B, et al. Fiber-coupling technique with micro step-mirrors for high-power diode laser bars [J]. Fiber Optic Components, 1997, 51(10): 390–393.
24. Anthony J, Horton, et al. Coupling light into few-mode optical fibres I: The diffraction limit [J]. Optics express, 2007, 15(4): 1443–1453.
25. Don M. Boroson, Bryan S. Robinson. The lunar laser communication demonstration: NASA's first step toward very high data rate support of science and exploration missions [J]. Space Science Reviews, 2014, 185(1-4): 115–129.
26. Robinson B S, Boroson D M, Burianek D A, et al. The NASA lunar lase communication demonstration-successful high-rate laser communications to and from the moon [C]//SpaceOps 2014 Conference. Pasadena, California, United States, 5–9 May, 2014: 1–8.
27. Carl M, Weinert M. Taper Device for Tolerant Coupling of Free-Space Optical Beams Into Single-Mode Fiber [J]. IEEE Photonics Technology Letters, 2016.
28. Jinsong Xiang, Yan Chen, Yu Hu. Effects of atmospheric turbulence on spatial optical coupling to single-mode fibers [J]. Intense Lasers and Particle Beams, 2006, 18(3): 377–380.
29. Ma Jing, Zhao Fang, Tan Li-Ying, et al. Plane wave coupling into single-mode fiber in the presence of random angular jitter [J]. Applied Optics, 2009, 48(27): 5184–51810.
30. Liu YT, Li B. Analysis of the effect of spot position jitter on spatial optical to single-mode fiber coupling efficiency [J]. Science and Technology Information, 2018, 16(25): 56–510.
31. Song Jiaxue, Chen Chunyi, et al. Study on the probability distribution of coupling efficiency of turbulence-disturbed single-mode fibers [J]. Advances in Lasers and Optoelectronics, 2021, 58(19): 140–148.
32. Wu Han-Ling, Yan Hai-Xing, Li Xin-Yang. Modal correction for fiber-coupling efficiency in free-space optical communication systems through atmospheric turbulence [J]. Optik - International Journal for Light and Electron Optics, 2010, 121(19): 1789–1793.
33. Han LQ, Wang Q. Fiber coupling efficiency and compensation for free-space optical communication under turbulent flow [J]. Optical Engineering, 2011, 38(05): 99–102.
34. Yang Q. B. Research on improving the coupling efficiency of spatial optical communication systems [D]. Harbin: Harbin Institute of Technology, 2013.
35. Xiong Jun, Ai Yong, Shan Xin, et al. Spatial optical communication fiber coupling efficiency and compensation analysis [J]. Infrared and Laser Engineering, 2013, 42(09): 2510–2516.
36. Luo W, Geng C, Li XY. Analysis and experimental study of the effect of atmospheric turbulence aberration on single-mode fiber coupling efficiency [J]. Journal of Optics, 2014, 34(06): 46–52.
37. Li F, Geng C, Li XY, et al. Research on adaptive fiber coupler array technology based on SPGD algorithm [J]. Infrared and Laser Engineering, 2015, 44(07): 2156–2161.
38. QI Xiaoli. Research on precise control and selection method of optical modes using spatial light modulator in air-division multiplexing [D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
39. Tu Jiajing, Zhang Huan, Li Han, et al. Design of multicore fiber-based triple-mode multiplexer/demultiplexer [J]. Journal of Optics, 2017, 37(03): 162–1610.
40. Shen Dong-Ya, Wang Chang-Hui, Machuan, et al. A novel optical waveguide LP01/LP02 mode converter [J]. Optics Communications, 2018, 4(18): 98–105.
41. Gao Hao, Yang Huajun, Xiang Jinsong. A method for automatic coupling of spatial optical-single-mode fibers [J]. Optical Engineering, 2007, 34(08): 126–1210.
42. Wu ZK, Chen Mo, Liu Chao, et al. A single-mode fiber coupling method based on grating helical scanning and SPGD algorithm [J]. Advances in Lasers and Optoelectronics, 2017, 54(06): 82–90.
43. Zhao Baiqiu, Meng Lixin, Yu Xiaonan, et al. Research on spatial optical to single-mode fiber chapter dynamic coupling technology [J]. China Laser, 2019, 46(11): 246–254.
44. Li Bo, Liu Yu-Tong, Tong Shou-Feng, et al. Adaptive Single-Mode Fiber Coupling Method Based on Coarse-Fine Laser Nutation [J]. IEEE Photonics Journal, 2018, 10(6): 1–13.

45. QI Yuanjing, Yan Ke, Zhou Xiongfeng. Research on accurate alignment algorithm of photodetector and single-mode fiber [J]. Manufacturing Automation, 2019, 41(12): 36–310.
46. Shi P, Li S, Zhang Guifen, et al. Microchip prism stack beam shaping and fiber-coupled output of high-power laser diodes [J]. Journal of Optics, 2000, 20(11): 1544–1548.
47. Bo JX, Qu Y, Gao X, et al. Fiber-coupled output of high-power array semiconductor lasers [J]. Photonics-Laser, 2001, 12(05): 468–470.
48. Zhou Chongxi, Liu Yinhui, Xie Weimin, et al. High power semiconductor laser array beam fiber coupling study [J]. China Laser, 2004, 31(11): 1296–1300.
49. Wang X.P., Liang X.M., Li Z.J., et al. 880nm semiconductor laser array and fiber coupling module [J]. Optical Precision Engineering, 2010, 18(05): 1021–1028.
50. Chen HN, Zou YG, Xu L, et al. Fiber coupling technology for high-power single-tube semiconductor lasers [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2014, 37(01): 6–10.
51. Xu D, Huang Xuesong, Jiang Menghua, et al. Development of high power and high efficiency fiber-coupled semiconductor laser module [J]. Intense Lasers and Particle Beams, 2015, 27(06): 69–74.
52. Liu Chang, Beiguang. ZEMAX-based fiber coupling design for multi-beam semiconductor lasers [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2015, 38(05): 22–25.
53. Liu S. W., Ren H., Wang W. Spatial coupling technology for high-power semiconductor lasers [J]. Electronic Fabrication, 2017, 23(13): 58–60.
54. Wang Xiaowei, Xiao Jianwei, Ma Snapdragon, et al. Fiber coupling of laser diode line array and multimode fiber array [J]. Journal of Semiconductors, 2002, 23(05): 464–468.
55. Lu D, Feng D, Zhang GW, et al. Fiber coupling of high-power semiconductor lasers [J]. China Laser, 2002: 29(s1): 21-28.
56. Wei R, Wang Yuzhu. Telescope collimation system for single-mode fiber coupling of lasers [J]. China Laser, 2003, 30(08): 687–690.
57. Zhao FY, Zhang Q, Tang HQ. Coupling of flat-ended fibers with tapered spherical lens fibers [J]. Journal of Photonics, 2003, 32(02): 218–221.
58. Liu YY, Yang RX, et al. Fiber coupling study of 974 nm semiconductor laser [J]. China Laser, 2014, 41(11): 23–29.
59. Shi K-R, Zhu C-Q. Design of combined lens system for semiconductor laser coupled with fiber [J]. Optical Technology Applications, 2016, 31(03): 17–110.
60. Li Shuquan. A method to improve the coupling efficiency of ball-end fiber [J]. Semiconductor optoelectronics, 1981, 20(03): 17–110.
61. Feng Tao. Research on a new optical antenna coupling system [D]. Wuhan: Huazhong University of Science and Technology, 2009.
62. Wang ZY, Qiu RH, Mo HT. A miniaturized optical antenna design for smoothing atmospheric turbulence effect [J]. Modern Telecommunication Technology, 2011, 41(7): 37-41.
63. Hu X, Zhang W-P, Yin R-G, et al. Coupling efficiency and transmission mode of laser in tapered multimode fiber [J]. Infrared and Laser Engineering, 2013, 42(02): 372–375.
64. Zhang SQ, Zhang Z, Cai L, et al. Single-lens-based spatial optical-single-mode fiber coupling method [J]. Strong Lasers and Particle Beams, 2014, 26(03): 41–45.
65. Yan, B. L., Li, X. L., Zhang, H. W., et al. Application of tapered optical fibers in spatial optical communication coupling systems [J]. Optical Precision Engineering, 2019, 27(02): 287–294.
66. Ren Lanxu, Zhang Jue, Xue Jingjing, et al. Analysis of the effect of temperature field on the coupling efficiency of spatial optical to single-mode fiber [J]. Space Electronics Technology, 2021, 18(05): 88–93.
67. Kang Jian-Mei, Guo Pan, Zhang Yin Chao, et al. Analysis of optimum coupling efficiency between random light and single-mode fiber [C]//ISPDI 2013-Fifth International Symposium on Photoelectronic Detection and Imaging. Beijing, China, 29 August, 2013: 277–285.
68. Luo C. H. Study on optical transmission characteristics of spatial optical-fiber coupled system [D]. Chengdu: University of Electronic Science and Technology, 2014.

69. Ke Xizheng, Lei Sichen, Wang Tianyu. Fiber-optic array-based Cassegrain transceiver antenna [J]. Advances in Laser and Optoelectronics, 2015, 52(12): 49–55.
70. Ke Xizheng, Yin Zhiyun. Coding theory in wireless laser communication systems [M]. Beijing: Science Press, 2016.
71. Kong Y. S., Ke X. Z., Yang Y.. Effect of local oscillation optical power on signal-to-noise ratio in spatially coherent optical communication [J]. Infrared and Laser Engineering, 2016, 45(02): 242–248.
72. Chen JN, Ke XZ. BER performance study of subcarrier-quadrature frequency division multiplexing system based on subcarrier outlier detection [J]. Journal of Optics, 2016, 36(02): 32–40.
73. Ke Xi-Zheng, Zhang Dan-Yu. Fuzzy control algorithm for adaptive optical system [J]. Applied Optics, 2019, 58(36): 9967–9975.
74. Chen JN, Ke XZ. Analysis of spatial optical-multimode fiber coupling efficiency under weak turbulence [J]. Journal of Xi'an University of Technology, 2014, 30(3): 326–330.
75. Lei Si-Chen. Research on optical coupling and beam control technology in free-space optical communication [D]. Xi'an: Xi'an University of Technology, 2016.
76. Zhang XT. Research on mode conversion method to improve coupling efficiency of single-mode optical fiber [D]. Xi'an: Xi'an University of Technology, 2020.
77. Yang Shang-Jun, Ke Xi-Zheng, Wu Jia-Li, et al. Dual-mirror adaptive-optics fiber coupling for free-space coherent optical communication [J]. Optical Engineering, 2021, 60(7): 1–13.
78. Yin Benkang. Experimental study of automatic alignment and control algorithms for spatial optical-fiber coupling [D]. Xi'an: Xi'an University of Technology, 2021.
79. Ke Xi-Zheng, Yin Ben-Kang. Experimental research on automatic alignment and control algorithm of spatial light-fiber coupling [J]. International Journal of Optics, 2021, 2021: 8481146
80. G. L. Wu, H. Z. Ke. Multi-dimensional coupler correction of spatial optical-single-mode fiber coupling alignment errors [J]. Journal of Optics, 2022, 42(7): 1–12.
81. Ke Xizheng, Wu Jali. Principles and applications of wireless optical coherent communication [M]. Beijing: Science Press, 2019: 20110.
82. Lei S. C., Ke X. Z.. Study of spatial optical coupling efficiency of lens arrays in atmospheric turbulence [J]. China Laser, 2015, 42(06): 179–186.
83. Ke Xizheng, Luo Jing, Lei Sichen. Automatic spatial optical coupling alignment method and implementation [J]. Infrared and Laser Engineering, 2018, 47(01): 93–910.
84. Ke XZ, Zhang XT. A method for improving single-mode fiber coupling efficiency using mode conversion: CN110133803A [P]. 2019-08-16.
85. Ke Xi-Zheng, Zhang Xu-Tong. Conversion of free-space optical path mode by spatial light modulator [J]. Optical Engineering, 2020, 59(01): 0161010.
86. Witham C R, Beranek M W, Carlisle B R, et al. Fiber-optic pigtail assembly and attachment alignment shift using a low-cost robotic platform [C]//50th Electronic Components and Technology Conference. Las Vegas, Nevada, United States, 21–24 May, 2000: 21–25.
87. Chen S.T., Yang D.X., Fang P.S., et al. Automatic coupling system for single-mode quartz fiber and plastic fiber [J]. Semiconductor optoelectronics, 2006(03): 282–285.
88. Thompson C A, Kartz M W, Flath L M, et al. Free space optical communications utilizing MEMS adaptive optics correction [C]//Free-Space Laser Communication and Laser Imaging II. Seattle, Washington, United States, 9 December, 2002: 129–139.
89. Cao Jing-Tai, Zhao Xiao-Hui, Li Zhao-Kun, et al. Stochastic parallel gradient descent laser beam control algorithm for atmospheric compensation in free space optical communication [J]. Optik, 2014, 125(20): 6142–6148.
90. Jiang Lun, Dai Zheng-Shuang, Yu Xin, et al. Experimental demonstration of a single-mode fiber coupling over a 1 km urban path with adaptive optics [J]. Journal of Russian Laser Research, 2021, 42(3): 363–370.



Atmospheric-Turbulence Models

14

Contents

Introduction 442
Atmospheric Refractive-Index Power Spectra 442
Turbulence Formation 442
Atmospheric Refractive-Index Fluctuation 444
Atmospheric Structure Constants 445
Atmospheric Turbulence Effects 448
Spectral Models of Atmospheric Turbulence 453
Foreign Research Progress 453
Domestic Research Progress 456
Power-Spectrum Models of Atmospheric Turbulence 459
Kolmogorov Spectrum 459
Tatarskii Spectrum 462
Von Kármán Spectrum 464
Greenwood–Tarazano Spectrum 464
Hill Spectrum 465
Andrews Spectrum 466
Non-Kolmogorov Spectrum 467
Turbulence Spectra Varying with Altitude 470
Atmospheric-Turbulence Channel Models 472
Log-Normal Distribution 472
Gamma–Gamma Distribution 474
Negative-Exponential Distribution 475
K Distribution 477
Malaga Distribution 478
Ke's Model 479
Summary 480
References 482

Abstract

In order to overcome the influence of atmospheric-turbulence effect on the performance of optical wireless communication system, it is necessary to establish a turbulence model which can accurately describe the fluctuations of

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_14

atmospheric refractive index. This paper summarizes the research progress of atmospheric-turbulence models at home and abroad, analyzes the theoretical establishment of different atmospheric-turbulence models in depth, and compares their application ranges. At the same time, the channel model of atmospheric established in the actual atmospheric-turbulence measurement is introduced. Finally, the development trend of atmospheric-turbulence model is forecasted.

Keywords

Optical wireless communication · Atmospheric-turbulence effect · Atmospheric-turbulence model · Kolmogorov spectrum · Atmospheric channel model

Introduction

In optical wireless communication links, atmospheric turbulence affects the optical signal at the receiving end, resulting in amplitude flicker, coherent attenuation, beam spread, and drift, which greatly limit the practical application of optical wireless communication. Therefore, a reasonable description of the refractive-index fluctuation of atmospheric turbulence has become key to the study of atmospheric-turbulence effects.

Atmospheric turbulence is a type of random motion. The turbulence-field parameters change irregularly with time and space. Random fluctuations of the atmospheric refractive index produce various atmospheric-turbulence effects, which seriously restrict the performance of optical wireless communication systems. To suppress the effects of atmospheric turbulence on the performance of optical wireless communication systems, it is necessary to establish a turbulence-spectrum model that can accurately describe atmospheric refractive-index fluctuations.

This chapter summarizes the research progress of atmospheric-turbulence spectral models by scholars at home and abroad in recent years, analyzes the theoretical establishment of different atmospheric-turbulence spectral models (Kolmogorov, Tatarskii, von Kármán, non-Kolmogorov, etc.), and compares their application ranges. In addition, it introduces the turbulence channel models established by experimental measurements in actual atmospheric turbulence, and finally forecasts the development trends of atmospheric-turbulence models.

Atmospheric Refractive-Index Power Spectra

Turbulence Formation

Following theoretical analyses, scholars have successively proposed a variety of atmospheric-turbulence spectral functions [1]. At present, most research on atmospheric-turbulence spectral models is based on the local homogeneous isotropy

of atmospheric turbulence (i.e., Kolmogorov statistical theory of turbulence). However, a large number of observational experiments in recent years have shown that most actual atmospheric turbulence tends to be anisotropic [2]. Therefore, establishing a more accurate atmospheric-turbulence spectral model is key for future research on statistical turbulence characteristics.

As shown in Fig. 14.1, turbulence was initially defined as irregular movement in a fluid, which is reflected in the formation of clouds in the atmosphere and movement of water in rivers. Daily life contains many turbulent flows, such as the blood flow in our own veins and arteries, and the flow of natural gas, oil, and water from their sources to the city interior.

Although turbulence is irregular, it has the following three characteristics [3]:

1. Irregularity and unpredictability: Turbulence is irregular in space and time, with unpredictable random patterns.
2. Statistical characteristics: Some statistical order appears in the irregularity of turbulent motion. The average and correlation are regular and predictable.
3. Mixing and enhanced diffusion: The fluid passes through complex and circuitous paths, resulting in a large amount of mixing of different parts of the fluid. This mixing significantly enhances the diffusion and increases the transfer of momentum, energy, heat, and other advections.

As shown in Fig. 14.2, Reynolds carefully observed the transition from laminar to turbulent flow in a series of experiments at the end of the nineteenth century [3]. First, still water in a water tank was slowly dropped into glass tubes of different diameters, the flow rate of the fluid in different tubes was changed, and the flow was observed in different tubes.

Reynolds believed that a dimensionless critical value could exist. Once the critical value was exceeded, the laminar flow would produce "bending" movement. The above experiments confirmed the existence of a critical value. Reynolds used the similarity theory to define this dimensionless quantity, called the Reynolds number [4]:

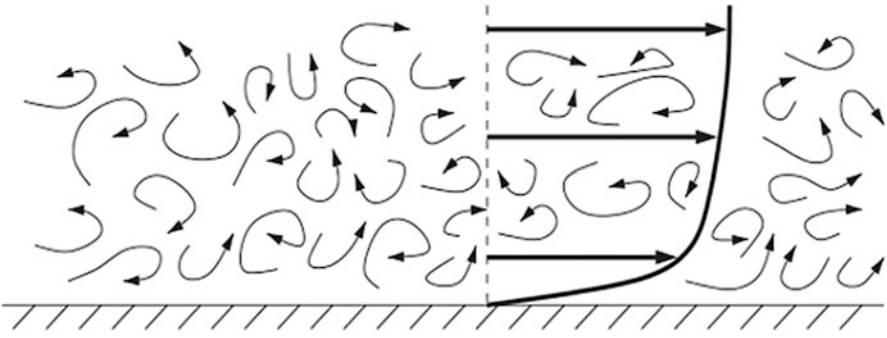


Fig. 14.1 Schematic diagram of irregular turbulent motion [3]

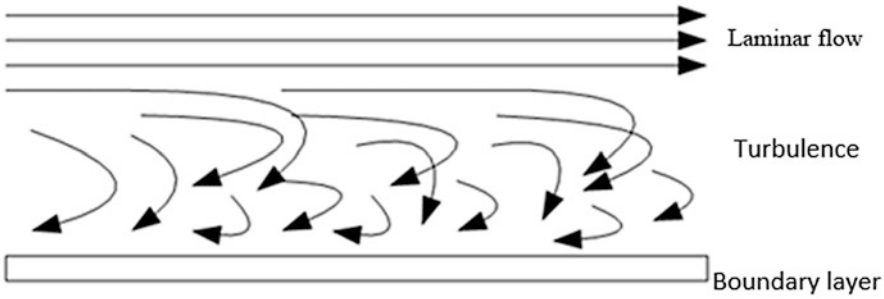


Fig. 14.2 Schematic diagram of laminar-flow and turbulent-flow transformation [3]

$$\mathrm{Re} = \frac{\rho \upsilon L}{\mu} = \frac{\upsilon L}{\nu}, \tag{14.1}$$

where L represents the fluid size (unit: m); υ represents the characteristic velocity (unit: m/s); and ρ indicates the fluid density. μ is the fluid's viscosity coefficient with dimension $MT^{-1}L^{-1}$ and $\nu = \frac{\mu}{\rho}$ is the kinematic viscosity coefficient with dimension M^2T^{-1}.

For smaller Re values, the viscosity dominates and the flow is laminar. For larger Re values, the inertial term dominates and the flow becomes more complex, eventually forming turbulence. In practical applications, different types of Reynolds numbers can be used according to the selection of the characteristic velocity and fluid size. In any case, the larger the Reynolds number, the more complex the flow.

Atmospheric Refractive-Index Fluctuation

The atmospheric refractive index is a random field that fluctuates slowly over time [2]. The random field of the atmospheric refractive index can be represented by $n(R, t)$, where t represents the time position, and $R = (x, y, z)$ represents the space position. The change in the atmospheric refractive index with time t is relatively slow, and the function $n(R)$ of the spatial position R is typically used to represent the sample function of the random field.

$\langle n(\mathrm{R})\rangle$ is the mean value of the sample function $n(R, t)$. When $\langle n(R)\rangle$ is independent of the spatial position R, it means that the statistics are uniform at this time. $B_n(R)$ is the covariance function of the random process, which is independent of the specific spatial location and is related only to the location interval $R = R_1 - R_2$. The expression for $B_n(R)$ is as follows [1]:

$$B_n(R) = \langle [n(R_1) - \langle n(R_1)\rangle][n^*(R_2) - \langle n^*(R_2)\rangle]\rangle, \tag{14.2}$$

where $\langle\rangle$ represents the statistical average, R_1 and R_2 are the positions of any two points in space, $n(R)$ is the function of the spatial position R, and $n^*(R)$ is the

conjugate complex function of $n(R)$. Generally, $n(R)$ is locally statistically uniform. The structure function $D_n(R)$ is typically used to describe the statistical characteristics of locally statistically uniform random fields [3] and is expressed as follows:

$$D_n(R) = \left\langle n(R_1 + R) - n(R_1)^2 \right\rangle, \tag{14.3}$$

where $D_n(R)$ is a structural function. Its relationship with the covariance function $B_n(R)$ is as follows:

$$D_n(R) = 2[B_n(0) - B_n(R)]. \tag{14.4}$$

Because $R = R_1 - R_2$, when $R_1 = R_2$, $B_n(R)$ becomes $B_n(0)$.

The power-spectral density $\Phi_n(K)$ and covariance function $B_n(R)$ form a Fourier transform pair; thus, they contain the same information but in different forms. The relationships are expressed as follows:

$$B_n(R) = \int e^{-iKR} \Phi_n(K) dK \tag{14.5}$$

$$\Phi_n(K) = \left(\frac{1}{2\pi}\right)^3 \iiint_{-\infty}^{\infty} e^{-iKR} B_n(R) d^3R, \tag{14.6}$$

where $K = (\kappa_x, \kappa_y, \kappa_z)$ represents the vector wavenumber, unit: *rad/m*; the module $\kappa = |K|$ represents the vector wavenumber, unit: *rad/m*. κ is also called the spatial wavenumber; it is the reciprocal of the spatial scale l; that is, $\kappa \sim 1/l$. e^{-iKR} is a complex variable function. Because the vector distance of $B_n(R)$ and $D_n(R)$ can be simplified into the scalar distance $R = |R_1 - R_2|$, the result can be expressed as

$$\Phi_n(\kappa) = \frac{1}{2\pi^2 \kappa} \int_0^{\infty} B_n(R) \sin(\kappa R) R dR. \tag{14.7}$$

The relationship $\Phi_n(\kappa)$ with $D_n(R)$ is derived:

$$\Phi_n(\kappa) = \frac{1}{4\pi^2 \kappa^2} \int_0^{\infty} \frac{\sin(\kappa R)}{\kappa R} \frac{d}{dR} \left[R^2 \frac{d}{dR} D_n(R) \right] dR. \tag{14.8}$$

Atmospheric Structure Constants

Atmospheric turbulence is anisotropic; however, it can be approximately regarded as a local homogeneous isotropy in a sufficiently small area. On this premise, the atmospheric refractive-index distribution meets the "two-thirds law" of the atmospheric refractive-index structure function $D_n(r)$ proposed by Kolmogorov [6]:

$$D_n(r) = C_n^2 r^{\frac{2}{3}} \quad (l_0 << r << L_0), \tag{14.9}$$

where r is the distance between two points of statistical turbulence characteristics, and l_0 and L_0 are the internal and external turbulence scales, respectively.

When the studied spatial scale is smaller than the inner turbulent scale l_0, viscous dissipation plays a major role in the influence of atmospheric fluctuation and inhibits the further development of turbulence. When the spatial scale is larger than the outer turbulent scale L_0, the inertial force of the air flow plays a major role in the atmospheric-fluctuation characteristics.

The atmospheric refractive-index structure constant C_n^2 is one of the basic parameters in atmospheric optics. Because the atmospheric structure constant on a slanted path is uneven, its size changes with the height. For a horizontal transmission, a typical value of the atmospheric structure constant is generally used. Therefore, to study the propagation characteristics of light beams in atmospheric turbulence, the atmospheric structure constant must be found.

Atmospheric turbulence can be divided into two categories: boundary layer and free atmosphere. The "boundary-layer" turbulence state is greatly affected by ground conditions; its thickness is about tens to hundreds of meters. The other is "free atmosphere" turbulence, which is basically unaffected by the ground. The boundary between the two layers is affected by the geographical environment, macrometeorological conditions, ground conditions, and other factors.

Boundary-layer turbulence is mainly affected by the surface temperature, humidity, and pressure. Wyngaard et al. proved that the relationship between the atmospheric structure constant in the boundary layer and the height is $-3/4$ power, which is a semi-empirical theory; however, the $-3/4$ power relationship well describes the change rule of the atmospheric structure constant over 100 m above the ground [7].

Later, Neff et al. also verified the $-3/4$ power relationship in Wyngaard's theory through an experiment of the change in C_n^2 with height measured by acoustic radar [8]. Kukharets and Tsvang later proposed an exponential model to describe the atmospheric refractive-index structure constant [9].

The atmospheric refractive-index structure constant C_n^2 fluctuates significantly over the course of 1 day, in addition to its variation with height. Figure 14.3 shows an example of the 24-h C_n^2 characteristics of the Rocky Mountains near Boulder, Colorado, at a distance of 2 m from the ground.

Ochs measured the change in the atmospheric structure constant with height h below a 3.2-km altitude. Hufnagel et al. provided a C_n^2 empirical formula applicable in the range of 3–24 km, based on measured data [11]:

$$C_n^2(h) = \left[2.2 \times 10^{-53} h^{10} (W/27)^2 \mathrm{e}^{-h/1000} + 10^{-16} \mathrm{e}^{-h/1500}\right] \cdot \mathrm{e}^{J(h,t)}. \tag{14.10}$$

In Eq. (14.10), J is a uniform Gaussian random variable with a zero mean value, and its correlation function is [12]

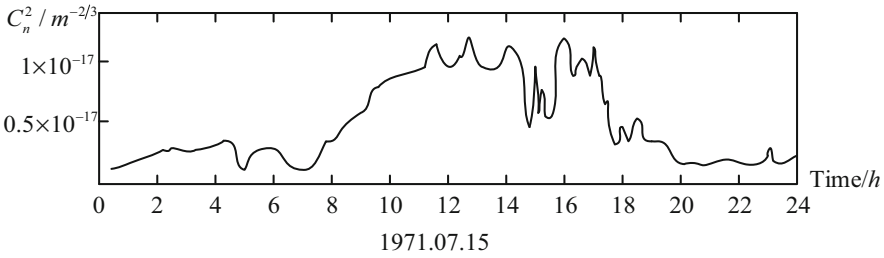


Fig. 14.3 24-h characteristics of C_n^2 on sunny days [10]

$$\langle J(h+h_1, t+\tau)\rangle = A\left(\frac{h_1}{100}\right)\mathrm{e}^{-\tau/5} + A\left(\frac{h_1}{100}\right)\mathrm{e}^{\tau/80}, \tag{14.11}$$

where

$$A\left(\frac{h_1}{100}\right) = \begin{cases} 1 - h/100 & h < 100 \\ 0 & h \geq 100 \end{cases}. \tag{14.12}$$

In Eqs. (14.11) and (14.12), h represents the height and τ is the time interval. The function W represents the root-mean-square velocity of the altitude between 5 and 20 km, expressed as

$$W = \left[\frac{1}{15}\int_5^{20} v^2(h)\mathrm{d}h\right]^{1/2}, \tag{14.13}$$

where (h) is the wind speed at the given altitude.

The typical data for W proposed by Hufnagel [13] is that the mean value is 27 m/s, and the standard deviation is 9 m/s. In 2001, the International Telecommunication Union-Radiocommunication Sector (ITU-R) issued a new atmospheric-structure model. This model makes it possible to theoretically analyze the oblique transmission of light waves [14]. Figure 14.4 shows the change in atmospheric structure constant with height issued by the ITU-R in 2001.

It can be observed that the atmospheric structure constant C_n^2 decreases with an increase in height. At 1–4 km, the atmospheric structure constant is unaffected by the wind speed with $C_n^2(0)$ at the surface. Above 10 km, it is affected by the wind speed [15]:

$$\begin{aligned} C_n^2(h) = {} & 8.148 \times 10^{-56} v_{RMS}^2 h^{10} \mathrm{e}^{-h/1000} \\ & + 2.7 \times 10^{-16} \mathrm{e}^{-h/1500} + C_0 \mathrm{e}^{-h/100}. \end{aligned} \tag{14.14}$$

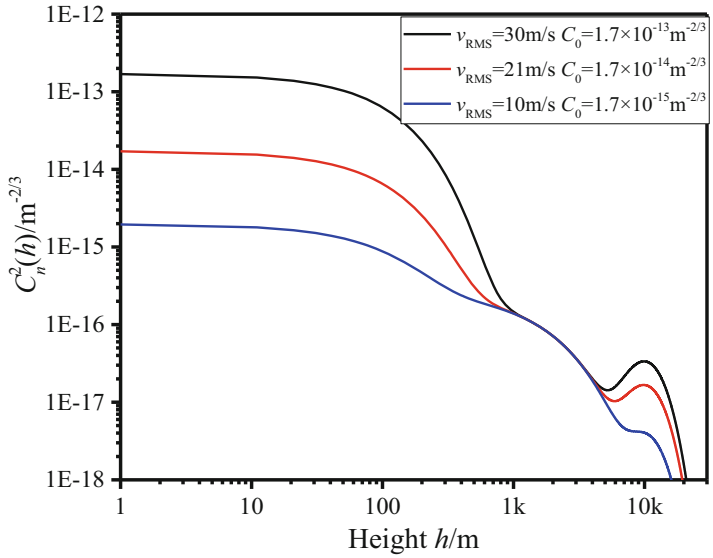


Fig. 14.4 Variation of the atmospheric structure constant with height h

Atmospheric Turbulence Effects

Beam Divergence

As shown in Fig. 14.5, F_0 is the curvature radius of the wavefront of the transmission plane. $F_0 = \infty$, $F_0 > 0$, and $F_0 < 0$ correspond to collimated, focused, and divergent beams, respectively. In the experiment, the beam is typically collimated or focused.

As shown in Fig. 14.6, in an optical wireless communication system, when the laser beam is transmitted in an ideal state, it is focused on the surface of the photodetector through a coupling lens, and then photoelectrically converted to obtain an electrical signal. The ideal facula obtained in the ideal state is denoted as Facula 1 in the figure (Fig. 14.7).

When propagating in atmospheric turbulence, because of the fluctuation of the atmospheric refractive index, the beam expands, and the energy of the beam cross section is reduced. This results in part of the beam not reaching the surface of the photoelectric detector through the coupling lens, such as the beam represented by the dotted line. The light spot obtained on the receiving surface is shown in Fig. 14.8. Atmospheric turbulence increases the light-spot radius and lowers the beam energy [16].

Beam Drift

As shown in Fig. 14.8, a wireless laser-communication system is composed of a transmitting antenna and a receiving antenna. Transmitting and receiving antennas

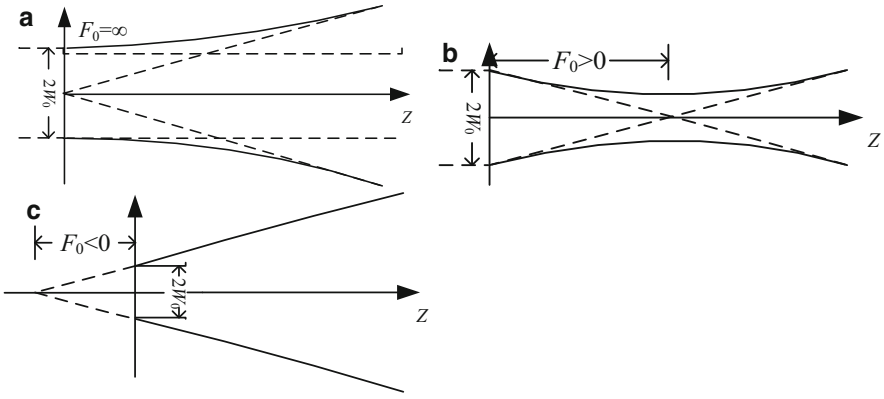


Fig. 14.5 Schematic diagram of different beams. (**a**) Collimated beam; (**b**) focused beam; (**c**) divergent beam

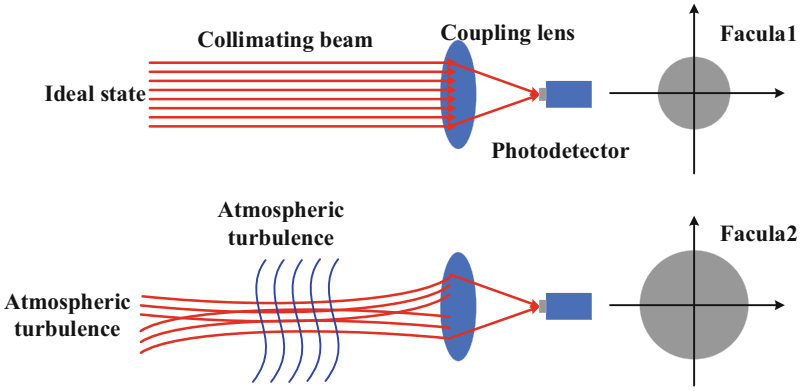


Fig. 14.6 Impact of beam spreading on atmospheric laser-communication system

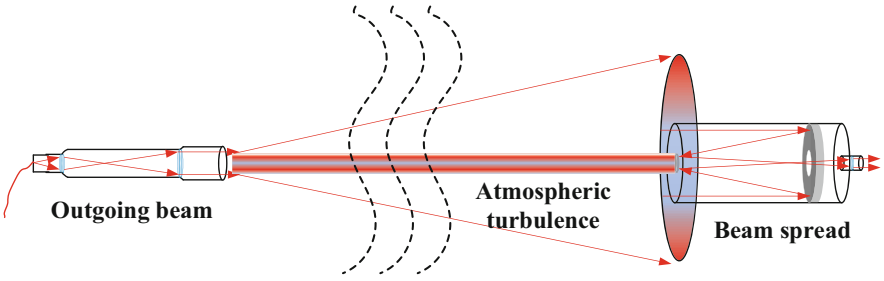


Fig. 14.7 Schematic diagram of beam spread

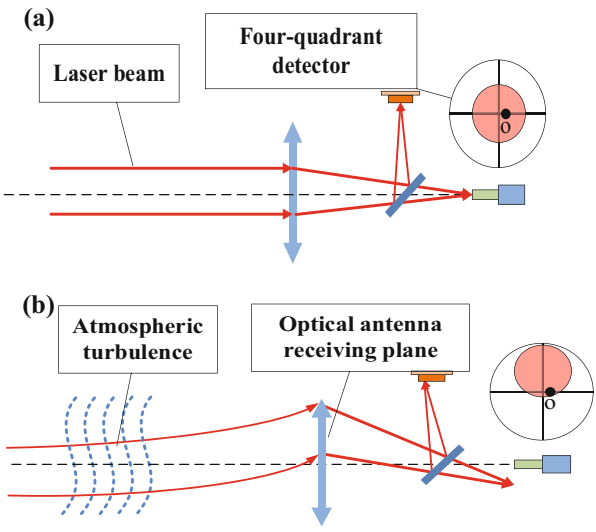


Fig. 14.8 Effect of spot drift on optical wireless communication system. (**a**) Ignoring atmospheric turbulence; (**b**) considering atmospheric turbulence

usually use a Cassegrain telescope; its optical structure can be simplified as a single-lens system. The following can be seen from Fig. 14.8:

1. When atmospheric turbulence is ignored, the transmitting beam is transmitted to the receiving surface of the receiving optical antenna, and its main optical axis coincides with the viewing axis of the receiving antenna. The beam focus is concentrated on the receiving end of the communication detector through the receiving optical system.

A beam-splitter prism is placed on the optical path inside the optical system to divide the receiving optical path into two parts: one for communication detection and the other for spot-position detection. A four-quadrant detector is used as the detection sensor. Using the spot position on the receiving surface of the four-quadrant detector, the center position of the spot can be obtained, and the spot center coincides with the center of the four-quadrant detector [17].

2. When atmospheric turbulence is considered, the beam drifts, owing to the influence of atmospheric turbulence during transmission. The beam moves randomly perpendicular to the optical axis and deviates from the receiving surface. At this time, the optical axis of the transmitted beam no longer coincides with the visual axis of the receiving antenna and includes an angle, which causes the focus of the beam converging through the optical antenna to shift. The shifted focus is no longer incident on the receiving surface of the communication detector.

At this time, the beam power received by the communication detector is significantly reduced, which adversely affects the signal detection of the communication system. The offset of the spot centroid, relative to the central position of the four-quadrant detector, can be detected through four-quadrant detection. With further turbulence development, the transmitting beam may drift completely out of the receiving surface of the receiving antenna. At this time, the receiving antenna cannot converge the transmitting beam, which interrupts the communication link [18].

Beam Spreading

As shown in Fig. 14.9, if we briefly observe the receiving plane, we can see that the distance from the widened spot with radius r_1 to the origin r_2 is the beam-drift amount. If the observation is sufficiently long, we can see that the random walk of the spot will produce a large spot with a mean-square radius of r_3. The former is called short-term beam spreading and the latter is called long-term beam spreading [19, 20].

Beam spreading refers to a change in the spot radius or area on the receiving plane caused by the turbulence vortex. When light passes through a small-scale vortex, it produces a diffraction effect, leading to beam expansion; in contrast, when it passes through a large-scale vortex, it is refracted, causing the beam to drift [21–23].

Light-Intensity Flicker

As shown in Fig. 14.10, when the beam diameter is much larger than the turbulence scale, the beam section contains multiple turbulence vortices, each of which will independently scatter and diffract the beam, causing the optical density on the detector plane to change continuously in time and space.

If the light intensity is measured at a certain distance from the light source in the turbulent atmosphere, the light intensity I will change with time and fluctuate randomly around the average value $<I>$, which is called light-intensity flicker. The magnitude of the flicker is directly related to the strength of the turbulence, and the flicker effect will increase the bit-error rate of the communication system [24–27].

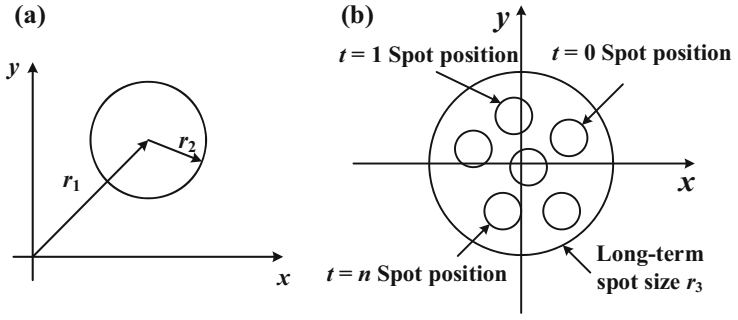


Fig. 14.9 Schematic diagram of a light spot observed on the receiving plane. (**a**) Short-term images; (**b**) long-term images

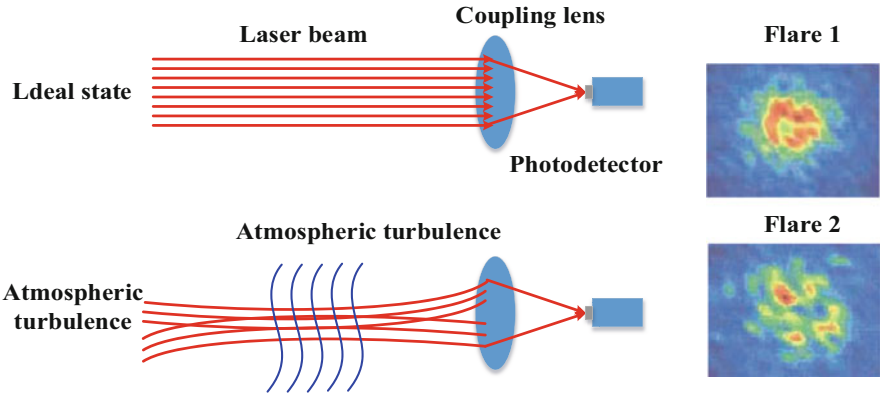


Fig. 14.10 Effect of light-intensity flicker on atmospheric laser-communication system

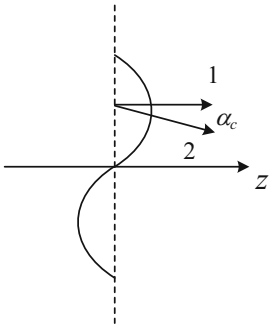


Fig. 14.11 Schematic diagram of the local arrival angle. (1) Uniform phase front normal; (2) equal phase normal

Angle-of-Arrival Fluctuation

As shown in Fig. 14.11, the laser has a uniform wavefront when propagating in a uniform medium. When propagating in a turbulent atmosphere, owing to the fluctuation of the atmospheric refractive index in different parts of the beam cross section, different parts of the beam wavefront undergo different phase changes [28, 29].

As shown in Fig. 14.12, atmospheric turbulence may cause laser-power attenuation, pulse broadening, intensity fluctuation, beam drift, angle-of-arrival fluctuation, and beam spreading. These decrease the signal-to-noise ratio of the communication system, increase the fluctuation variance, and increase the bit-error rate. Compared with fully coherent light, partially coherent light can suppress spot drift, intensity flicker, and angle-of-arrival fluctuation.

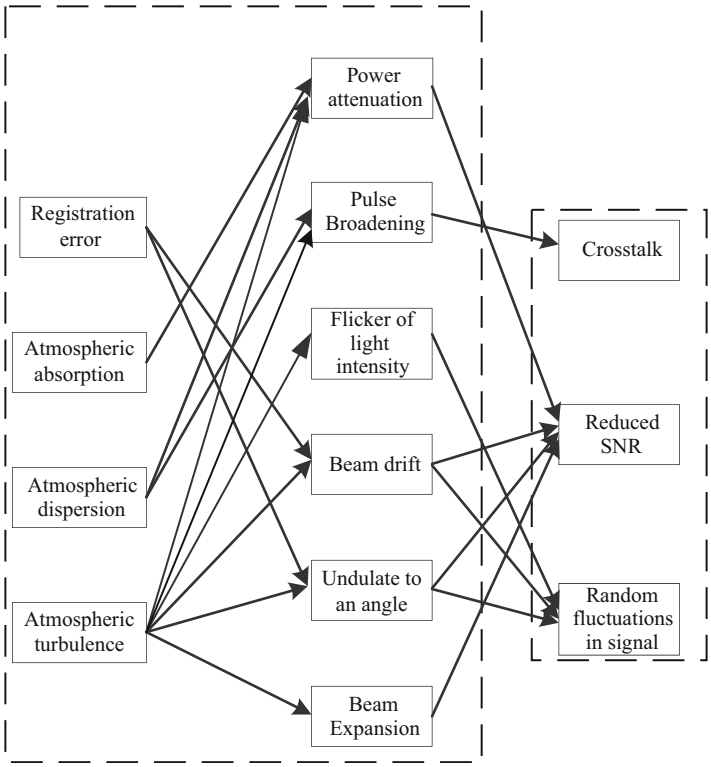


Fig. 14.12 Impact of atmospheric channels on optical communications

Spectral Models of Atmospheric Turbulence

Foreign Research Progress

In his famous turbulence theory [5] published in 1941, Kolmogorov formally proposed the concept of an atmospheric-turbulence power-spectrum model and established the Kolmogorov spectrum model. Subsequently, researchers worldwide began to investigate the atmospheric-turbulence power spectrum and successively established different atmospheric-turbulence spectrum models. The most commonly used models include the Kolmogorov [30], von Kármán [31], Hill [38], and non-Kolmogorov spectra [53].

In 1941, Kolmogorov et al. established the theory of locally isotropic turbulence and proposed that turbulence should be approximated as a locally isotropic Kolmogorov spectrum. In the inertial subregion, the relationship between the atmospheric refractive-index structure function and turbulence structure constant followed the well-known two-thirds law, and the power index of the power spectrum was $-5/3$ [30].

In 1948, von Kármán et al. proposed the von Kármán spectrum to describe the turbulence generated in a circular tube. This spectrum included the influence of both internal and external scales, and avoided the singularity of the power spectrum at the origin. This is applicable to the full wavenumber region [31].

In 1961, Tatarskii et al. established the Tatarskii spectrum by introducing an internal-scale influence factor as the truncation function and extended the application area of the spectrum to the dissipation area (i.e., the area where the energy dissipation exceeds the kinetic energy) [32].

In 1963, von Kármán et al. considered the influence of internal and external turbulence scales, based on the Kolmogorov spectrum, and established a modified von Kármán spectrum, which solved the problem that the Kolmogorov and Tatarskii spectra could not integrate singular points [33].

In 1968, Busch et al. experimentally measured the atmospheric turbulence in different regions. The results showed that the measured one-dimensional velocity spectrum usually followed the $-5/3$ power law, and the wavenumber was far less than that expected within the spectral-inertia subrange [34].

In 1972, Kaimal et al. conducted a study on the turbulence spectrum in a near-ground environment. The results showed that the turbulence power spectrum in the inertial subinterval satisfied Kolmogorov's $-5/3$ law and conformed to the isotropic property [35].

In 1974, Greenwood et al. experimentally measured the turbulence power spectrum on a rough underlying surface and proposed the Greenwood–Tarazano spectrum, based on experimental data [36].

In 1977, Champagne et al. conducted an experimental study on the turbulence spectrum in Minnesota, U.S. The study found that all power-spectral densities showed a bulge at high wavenumbers. When the wavenumbers gradually approached the high-wavenumber area, the spectrum deviated from the $-11/3$ rule and declined rapidly [37].

In 1978, Hill considered the influence of sudden changes in the power spectrum in the region close to a high wavenumber, and proposed a numerical model for accurately describing the spectrum in the inertial and dissipation regions (Hill spectrum) [38].

In 1986, Manning et al. proposed a model to describe the anisotropy of refractive-index fluctuations near the Earth's surface and derived the parameters of the model based on experimental data. The results showed that the anisotropy of the atmospheric-turbulence field could be verified by the difference in the transverse mutual-interference function of the propagating plane wave [39].

In 1990, Kraichnan et al. studied the Kolmogorov spectrum in the inertial range when Fourier mode, homogeneity, and local isotropy were not natural assumptions, gave effective scaling parameters, and generalized the $-5/3$ law for the eigenvalue spectrum of the two-point velocity correlation matrix [40].

In 1992, Andrews et al. proposed the Andrews spectral model based on the Hill spectral model, which reasonably described the protrusion of the power spectrum in the high-wavenumber region [41].

In 1995, based on the Kolmogorov turbulence spectrum, Beland expressed the structure function and turbulence spectrum describing non-Kolmogorov turbulence, and the formulas for the logarithmic-amplitude fluctuation variance and coherent length under weak fluctuation conditions [42].

In 1995, Emaleev et al. investigated the anisotropy of the turbulence spectrum of the ground atmosphere. The results showed that the anisotropy of the low-frequency characteristics of the turbulence spectrum depended largely on the height above the underlying surface [43].

In 1998, Stribling et al. experimentally observed that the exponential value of the measured power function was significantly different from the eigenvalue of the Kolmogorov theoretical spectrum. Based on the assumption of turbulence isotropy, they studied the basic relationship between the exponential structural function and exponential power spectrum [44].

In 1999, Mikhail et al. studied the anisotropy of the horizontal and vertical tilt components caused by the anisotropy of turbulence near an optical antenna. The results showed that, owing to the turbulence anisotropy, the power spectrum of the horizontal tilt exceeded that of the vertical tilt [45].

In 2004, Lipatack et al. defined the generalized Tatarskii spectrum, which expanded the structure function to a smaller power-law value, had singularity when $K = 0$, and was effective for all beams [46].

In 2007, Phillips et al. compared the von Kármán spectrum with the Hill spectrum and obtained a method to estimate the external atmospheric-turbulence scale, based on the correlation function of the transverse displacement of a horizontal transmission beam [47].

In 2008, Zilberman et al. studied the turbulence characteristics of a communication link under a change of the logarithmic amplitude, based on the three-layer model of the refractive-index fluctuation spectrum in a free atmosphere. They proposed a turbulence-spectrum model that varied with altitude, according to the experimental and theoretical data of the troposphere and the lower-stratosphere turbulence profile [48].

In 2008, Toselli et al. proposed a non-Kolmogorov turbulence power-spectrum model that explained the heterogeneity and anisotropy of atmospheric turbulence [49].

In 2010, Zilberman et al. analyzed the impact of Kolmogorov and non-Kolmogorov turbulence statistics on the laser-communication link for different transmission scenarios [50].

In 2014, Toselli et al. proposed an anisotropic atmospheric-turbulence spectrum and described in detail the asymmetry of atmospheric-turbulence vortices in the vertical and horizontal directions [51].

In 2018, Helin et al. observed the problem of joint estimation of a turbulence profile and unknown power-spectral density on the ground. They proposed three numerical-reconstruction methods that could accurately locate local disturbances under the non-Kolmogorov power-spectral density [52].

In 2021, Santos et al. studied nearly isotropic turbulence through experiments, determined the flow-field characteristics that affected the dissipation frequency, and modified the von Kármán spectrum to simulate the dissipation range [53].

In 2021, Goldman et al. used an analytical model of steady turbulence to obtain the inertial-range power spectrum of compressible turbulence. The results showed that the rate of controlling the energy transferred to all relatively large quantities through turbulent viscosity decreased, and the inertial range was a power law with exponent −2 [54].

In 2021, Hosoi et al. adopted the inertial-dissipation method of the Kolmogorov turbulence theory and von Kármán spectral model. The analysis results showed that the average value of the logarithm of the turbulent dissipation rate increased with decreasing height [55].

Domestic Research Progress

With the continuous investigation of the optical characteristics of atmospheric turbulence, domestic scholars have conducted a large number of experimental measurements and theoretical analyses on the atmospheric-turbulence power spectrum from different regions, meteorological conditions, and underlying surfaces. They further improved the theory of atmospheric-turbulence spectrum models and laid a solid foundation for subsequent research on atmospheric turbulence.

In 1975, Zhou et al. defined a vortex Reynolds number R_a, established the relationship between the vortex scale and vortex Rayleigh number, and derived the turbulent energy-spectrum exchange function related to the binary velocity [56].

In 1978, Liu et al. proposed a dimensionless quantity R_i to describe the characteristics of the atmospheric-turbulence power spectrum, based on the distribution of an inversed turbulence spectrum obtained from a large amount of measured atmospheric-turbulence power-spectrum data, and used this dimensionless quantity to divide the spectral area [57].

In 1984, An et al. proposed a formula for calculating the energy-spectrum density of homogeneous isotropic turbulence. The results showed that the turbulence-energy spectrum in the wavenumber range of the dissipation zone was consistent with the actual measurement results [58].

In 1986, Song et al. studied the influence of the von Kármán turbulence spectrum and exponential spectrum on beam drift in weak and strong fluctuation regions, respectively. The results showed that the theoretical values obtained with the different turbulence spectra were consistent with the experimental results [59].

In 1986, based on the frozen-fluctuation hypothesis, Wu et al. used the path-integral method of the parabolic-wave equation to define the relevant scale of the generalized refractive-index fluctuation power spectrum, in the case of a strong fluctuation limit with diffraction parameter λ. The results showed that when the propagation distance was far greater than the transverse scale of the refractive-index fluctuation, the refractive-index correlation scale had little influence on the optical-wave fluctuation characteristics [60].

In 1992, Wang et al. conducted turbulence-observation experiments in Lanzhou City, China, and analyzed the wind-speed component spectrum, temperature spectrum, and momentum co-spectrum. The results showed that the terrain features had little influence on the turbulence-spectrum features [61].

In 1993, Xu et al. studied the intensity, correlation coefficient, scale, and velocity spectrum of atmospheric-surface turbulence in Guangzhou, China. The results showed that in the urban near-surface layer, the turbulence behavior was consistent with the local isotropy law of the inertial subregion, which conformed to the Kolmogorov spectrum theory [62].

In 1998, Hu et al. used discrete orthogonal wavelets to conduct a multiscale decomposition of the boundary-layer turbulence fluctuation signals. The results showed that the energy spectral density was a power-exponential function of the frequency, and the turbulence spectral scale index met the $-5/3$ of the Kolmogorov spectrum [63].

In 1998, Zeng et al. analyzed the near-ground temperature power spectrum of a complex terrain in an actual atmosphere. The results showed that the power rate of the temperature spectrum was steep for complex terrain, and was between $-5/3$ and -3, which did not conform to the Kolmogorov turbulence-spectrum theory [64].

In 1999, Rao analyzed the turbulence intermittency and its model and established an intermittency model of the turbulent temperature field according to experimental data [65].

In 2000, Rao et al. proposed a theoretical method to analyze the temporal and spatial characteristics of non-Kolmogorov turbulent phase fluctuations and obtained their time-structure function and power spectrum [66]. In the same year, they proposed a new method to measure the power-spectrum exponential-decline factor for non-Kolmogorov turbulence [67].

In 2005, Zhu et al. proposed a method for simultaneously inverting the turbulence internal scale and refractive-index structure constant using the Andrews spectral model [68].

In 2005, Zhang et al. summarized the research status of the atmospheric-turbulence power spectrum and established a simplified turbulence power-spectrum model, based on the modified von Kármán spectrum [69].

In 2005, Liu et al. performed experimental measurements of the local turbulence isotropy. The results showed that in the inertial subregion, the turbulence spectrum conformed to the $-2/3$ exponential law of Kolmogorov turbulence; however, in the low-frequency region, it was in a decentralized state [70].

In 2006, Ma et al. measured the atmospheric-scintillation spectrum by building an 800-m laser-transmission experimental platform. The results showed that the high-frequency power index in the actual atmospheric-turbulence spectrum exhibited an unstable random fluctuation, in which the decline rate of the high-frequency spectrum was closely related to the distribution of the dissipation area of the turbulence spectrum [71].

In 2007, Gong et al. simulated the turbulence power-spectrum characteristics based on a numerical-prediction model. The slope of the power spectrum obtained

from the experimental simulation was nearly twice that of the turbulence power spectrum in the actual atmosphere [72].

In 2008, Wu et al. established a near-surface atmospheric estimation and prediction model and verified its validity using field measurement data [73].

In 2008, Wu et al. tested the hypothesis of local homogeneity on the surface of snow and fitted the spectral power rate to the normalized spectrum. The fitting result was 85%, which was consistent with −5/3 [74].

In 2010, Zhao et al. experimentally measured atmospheric-turbulence characteristics using microwave rangefinders and established a preliminary framework for a prediction model for sea-surface atmospheric refractive-index fluctuations [75].

In 2010, Ma et al. modified the non-Kolmogorov power spectrum based on the consistency between the refractive-index structure function and the power-spectral density and estimated the impact of non-Kolmogorov turbulence on the angle-of-arrival fluctuation [76].

In 2011, Jin et al. simulated multichannel radiant-brightness temperatures by using the atmospheric data measured during typhoon Phoenix. They proposed an improved atmospheric-absorption model and built a statistical matrix for retrieving atmospheric temperature profiles [77].

In 2011, Xue et al. deduced a generalized modified atmospheric-spectral model containing limited internal and external turbulence scales. The power-law value of this spectrum was between 3 and 5, instead of the standard power-law value of 3/12. When the internal and external scales were set to zero and infinity, respectively, the spectral model was simplified to a classic non-Kolmogorov spectrum [78].

In 2014, Li studied the change in the working frequency of the atmospheric-noise spectral density in the range of 10 Hz–1 kHz. The results showed that the atmospheric-noise level in autumn was significantly higher than that in spring, and the atmospheric-noise spectral density was close to a log-normal distribution [79].

In 2014, Ai et al. started with Kolmogorov's atmospheric-turbulence theory, elaborated on Kolmogorov's turbulence spectrum, and compared the similarities and differences between the Kolmogorov, Tatarskii, and other models [80].

In 2016, Li et al. obtained the anisotropic power spectrum of hypersonic-turbulence refractive-index fluctuation by transforming the generalized anisotropic von Kármán spectrum [81].

In 2018, Wang et al. conducted an experimental study on laser transmissions with a 610-m link length in Lanzhou. They measured the atmospheric refractive-index structure constant in sunny, overcast, dusty, sleet, and rainy weather conditions using the scintillation method. The results showed that atmospheric turbulence was medium intensity in all weather types [82].

In 2018, Zhang et al. built an indoor simulation device for non-Kolmogorov turbulence. They found that the power-law change in the time spectrum was primarily affected by the atmospheric-turbulence model [83].

In 2019, Changqiong Hu et al. introduced the calculation methods for various types of atmospheric-turbulence characteristic parameters and conducted experimental measurements of the atmospheric-turbulence power spectrum near the

ground level at the bridge location. The results showed that the power spectrum density increased by 1 to 2 orders of magnitude at high wind speeds compared to low wind speeds [84].

In 2021, Zhao et al. established a non-Kolmogorov turbulence spectrum with an anisotropic tilt angle on horizontal and satellite links, and derived the analytical expression of the scintillation index of the anisotropic non-Kolmogorov horizontal and satellite links [85].

In 2022, Yang et al. conducted an experimental study on the spectral characteristics and similarity laws of turbulence in the Qinghai, Tibet region. The results showed that the slope of the inertial subzone of the wind speed and temperature spectra in all directions followed the $-2/3$ rule, and both met the isotropic conditions [86].

Power-Spectrum Models of Atmospheric Turbulence

Because the turbulent-flow field is a random function of time and space variables, statistics are required to describe it. To fully understand and describe the statistical characteristics of a turbulent-flow field, it is necessary to understand the atmospheric refractive-index power spectrum. However, because it is difficult to accurately obtain the statistical properties of high-order moments theoretically and experimentally, turbulence statistical theory usually focuses on the statistical properties of low-order moments, especially the second-order statistical properties.

The power-spectrum model of atmospheric refractive-index fluctuation differs in different turbulent environments. To further study the optical-transmission characteristics in atmospheric turbulence, scholars have successively proposed a variety of atmospheric refractive-index power-spectrum models suitable for different conditions, as shown in Tables 14.1 and 14.2.

The atmospheric refractive-index power spectrum model was developed from the original Kolmogorov spectrum, with an assumed external scale $L_0 \to \infty$ and internal scale $l_0 = 0$, to the von Kármán spectrum, which reflects the characteristics of internal and external scales, to the more accurate universal Hill spectrum model, as well as the non-Kolmogorov spectrum, which is most consistent with the actual atmospheric characteristics at present [87]. The development of these models also marks the continuous improvement of current theoretical turbulence research.

Kolmogorov Spectrum

In 1941, Kolmogorov proposed three assumptions about the equilibrium state in uniform turbulence at a sufficiently high Reynolds number and sufficiently small scale, and thus established the basis of classic turbulence theory [30]:

Table 14.1 Power-spectrum models of the atmospheric-turbulence refractive index

Power-spectrum name	Function
Kolmogorov	$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3}$
Tatarskii	$\Phi_n(\kappa) = 0.33 C_n^2 \kappa^{-11/3} \exp\left(\frac{\kappa^2}{\kappa_m^2}\right)$
Von Kármán	$\Phi_n(\kappa) = 0.33 C_n^2 \kappa^{-11/3} \exp\left(\frac{\kappa^2}{\kappa_m^2}\right)$
Modified von Kármán	$\Phi_n(\kappa) = 0.33 C_n^2 \frac{\exp\left(-\kappa^2/\kappa_m^2\right)}{\left(\kappa^2+\kappa_0^2\right)^{11/6}}$
Greenwood–Tarazano	$\Phi_n(\kappa) = 0.33 C_n^2 \left(\kappa^2 + \kappa\kappa_0\right)^{-11/6}$
Hill	$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3} \times \left\{ \exp\left(-1.2\kappa^2 l_0^2\right) + 1.45 \exp\left[-0.97(\ln \kappa l_0 - 0.452)^2\right] \right\}$
Andrews	$\Phi_n(\kappa) = 0.033 C_n^2 \left[1 + a_1(\kappa/\kappa_l) - a_2(\kappa/\kappa_l)^{7/6}\right] \frac{\exp\left(-\kappa^2/\kappa_l^2\right)}{\left(\kappa_0^2+\kappa^2\right)^{11/6}}$
Non-Kolmogorov	$\Phi_n(\kappa, \alpha) = \Gamma(\alpha - 1)\cos(\alpha\pi/2)/(4\pi^2)\widetilde{C}^2{}_n \kappa^{-\infty}$
Generalized modified non-Kolmogorov	$\Phi_n(\kappa, \alpha) = \frac{A(\alpha)\widetilde{C}_n^2}{\left(\kappa^2+\kappa_0^2\right)^{\alpha/2}} \exp\left(-\kappa^2/\kappa_m^2\right)$

Table 14.2 Classification diagram of atmospheric-turbulence refractive-index power-spectrum models

Turbulence type	Power-spectrum type	Power spectrum	Internal-scale influence	External-scale influence
Kolmogorov turbulence (power-law value is 11/3)	Basic	Kolmogorov	Not considered	Not considered
	Mathematical derivation	Tatarskii	Considered	Not considered
		Von Kármán	Considered	Considered
	Physical analysis	Greenwood–Tarazano	Considered	Considered
		Hill	Considered	Considered
		Andrews	Considered	Considered
Non-Kolmogorov turbulence (power-law value is between 3 and 5)	Basic	Non-Kolmogorov	Not considered	Not considered
	Mathematical derivation	Generalized modified non-Kolmogorov	Considered	Considered

1. Small-scale turbulent motion is statistically isotropic.
2. The statistics of small-scale turbulent motion have the energy-dissipation rate ε and kinematic viscosity v, which is the only general form to be determined. This scale range is also called the universal equilibrium range.
3. Large-scale turbulent-motion statistics have a unique ε, whose general form can be determined and is independent of v.

Turbulence satisfying these three assumptions is called Kolmogorov turbulence [30]. The Kolmogorov theory of local isotropic turbulence allows for nonuniformity and anisotropy on a large scale. It is assumed that with the cascade transfer of energy to a small scale, the directional effect generated on a large scale becomes increasingly smaller and weaker, until the vortex motion becomes statistically uniform, isotropic, and independent of the specific energy-generation mechanism.

Based on the research results of Kolmogorov's turbulence theory, turbulence can be divided into input, inertia, and dissipation areas, according to the size of the spatial wavenumber κ (see Fig. 14.13). In the figure, r represents the turbulence scale, and l_0 and L_0 are the internal and external turbulence scales, respectively.

The region of $\kappa < 2\pi/L_0$ is the input region, in which $\Phi_n(\kappa)$ is usually anisotropic. The area of $2\pi/l_0 > \kappa > 2\pi/L_0$ is the inertial area; at this time, the shape of $\Phi_n(\kappa)$ is determined by the physical laws that restrict the breaking of large turbulent vortices into small ones. When κ continues to increase until $\kappa > 2\pi/l_0$, the turbulence area enters the dissipation area. The turbulent energy in the dissipation area dissipates much more than the generated kinetic energy, so $\Phi_n(\kappa)$ declines significantly more rapidly.

Kolmogorov proposed that when R is less than L_0, the turbulence is locally statistically homogeneous and isotropic [6]. In the inertial subregion, the

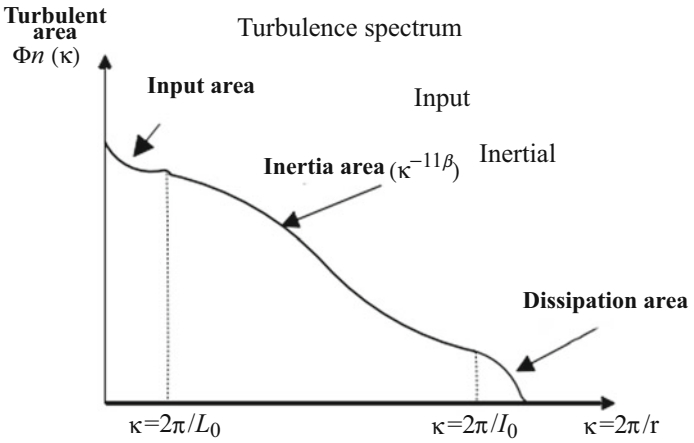


Fig. 14.13 Kolmogorov turbulence-spectrum model [4]

refractive-index structure function $D_n(R)$ is proportional to the two-thirds power of the scalar distance R:

$$D_n(R) = C_n^2 R^{2/3}, \tag{14.15}$$

where C_n^2 is the atmospheric refractive-index structure constant, and the unit is $m^{-2/3}$. The Kolmogorov spectrum can be obtained by substituting Eq. (14.15) into Eq. (14.16):

$$\begin{aligned} \Phi_n(\kappa) &= \frac{\Gamma(a+1)}{4\pi^2} \sin[(a-1)\pi/2] C_n^2 \kappa^{-a-2} \\ &= 0.033 C_n^2 \kappa^{-11/3}, 1/L_0 \le \kappa \le 1/l_0, \end{aligned} \tag{14.16}$$

where $a = 5/3$ and $\Gamma(a)$ represents the gamma function.

Tatarskii Spectrum

In the atmospheric-boundary layer, the characteristic-length scale of turbulent motion is the turbulence scale, defined as follows [5]:

$$L = \frac{1}{\sigma^2} \int_0^\infty R(r) dr, \tag{14.17}$$

where σ^2 is the variance of the velocity component or temperature, r is the spatial displacement, and $R(r)$ is the sum of the autocorrelation functions of the velocity component or temperature. L is the turbulence scale.

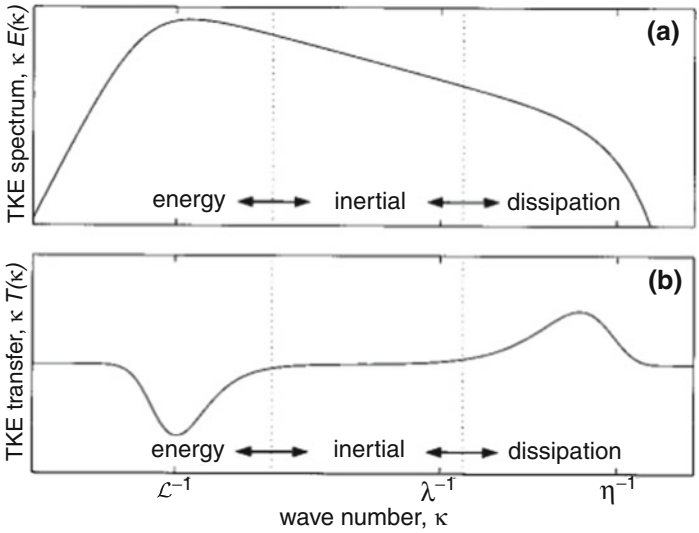


Fig. 14.14 Spectral map of the corresponding turbulence scale [5]

Figure 14.14a shows the spectrum of turbulent kinetic energy (TKE), where $E(k)$ is the TKE per unit mass. Figure 14.14b shows the TKE spectral conversion, where $T(k)$ is the amount of energy transferred from the wavenumber to all other wavenumbers.

Kolmogorov's inertial subrange scale is only applicable to the smaller turbulent scale. However, in actual atmospheric turbulence, there is no universal value for the turbulence scale, owing to the influence of atmospheric conditions, height above the ground, and terrain structure.

Tatarskii obtained a spectrum function by introducing a turbulence internal-scale parameter ($\kappa_m = 5.92/l_0$), which has the form of a Gaussian function, $\exp\left(\frac{\kappa^2}{\kappa_m^2}\right)$. For mathematical-description convenience, Tatarskii improved the power spectrum [15] used by Novikov in the study of turbulent wind-speed fluctuation, applied it to the study of atmospheric refractive-index fluctuation, and established the Tatarskii spectrum [33]:

$$\Phi_n(\kappa) = 0.33C_n^2\kappa^{-11/3}\exp\left(\frac{\kappa^2}{\kappa_m^2}\right), \kappa > 1/L_0, \kappa_m = 5.92/l_0. \tag{14.18}$$

The Tatarskii spectrum can only be applied to the dissipation area, and the definition domain of the Eqs. (14.16) and (14.18) spectral model has a singular point ($L_0 \to \infty$). This means that the structure function $D_n(R)$ in Eq. (14.15) can be calculated; however, the covariance function $B_n(R)$ in Eq. (14.14) cannot be calculated. Therefore, the Tatarskii spectrum is isotropic only when it is in the inertial region or outside the dissipative region.

Von Kármán Spectrum

The Kolmogorov and Tatarskii spectra are homogeneous and isotropic only outside the inertial region or dissipation region (i.e., $\kappa = 1/L_0$); when $\kappa < 1/L_0$, it is generally considered to be nonuniform, and its form is unknown [35]. The von Kármán spectrum improves the Kolmogorov and Tatarskii spectrum models and expands the range of κ to 0–∞; thus, it is also uniform and isotropic in the range of $\kappa < 1/L_0$. Therefore, the spectrum can contain all the beam ranges.

When the spatial wavenumbers, $D_n(R)$ and $B_n(R)$ of the von Kármán spectrum also exist simultaneously, which solves the problem that the Kolmogorov and Tatarskii spectra cannot integrate singular points. The von Kármán spectral expression is [7]

$$\Phi_n(\kappa) = 0.33C_n^2\left(\kappa^2 + \kappa_0^2\right)^{-11/6}, 0 \le \kappa \le \infty, \kappa_0 = 2\pi/L_0. \tag{14.19}$$

When the inertial region is $\kappa_0 \le \kappa \le \kappa_m$, Eq. (14.19) can be transformed into the Kolmogorov spectrum in Eq. (14.16).

To fully consider the influence of the inner and outer turbulence scales, the inner turbulence-scale parameter κ_m is also introduced into the von Kármán model, and a modified von Kármán spectrum is proposed. The spectrum expression is as follows [31]:

$$\Phi_n(\kappa) = 0.33C_n^2 \frac{\exp\left(-\kappa^2/\kappa_m^2\right)}{\left(\kappa^2 + \kappa_0^2\right)^{11/6}}, 0 \le \kappa \le \infty; \kappa_m = 5.92/l_0, \kappa_0 = 2\pi/L_0, \tag{14.20}$$

where $0 \le \kappa \le \infty$, $\kappa_0 = 2\pi/L_0$, and $\kappa_m = 5.92/l_0$. When the inertial region is $\kappa_0 \le \kappa \le \kappa_m$, Eq. (14.20) can be transformed into the Kolmogorov spectrum. However, outside the inertial region, Eq. (14.20) does not reflect the influence of mutation factors in the high-wavenumber region, so it cannot accurately describe the actual physical phenomena [40]. In view of their convenience in theoretical calculations, the Tatarskii and von Kármán spectra have become widely used models for subsequent research on refractive-index power-spectrum theory.

Greenwood–Tarazano Spectrum

The von Kármán spectrum describes the fluid in a pipe, which has only one characteristic dimension, namely, the pipe diameter. However, actual atmospheric turbulence has many characteristic scales because of the height, boundary-layer thickness, and nonuniform ground scales in the atmospheric boundary layer; therefore, the von Kármán spectrum cannot fully describe the actual atmospheric turbulence [39].

Because the temperature spatial-structure function is related to the spatial power spectrum and does not depend on the Taylor freezing hypothesis, Greenwood et al.

reconstructed the refractive-index spectrum by using a spatial structure function measured experimentally and compared it with the von Kármán spectrum. The results showed that the von Kármán spectrum is not the best model, according to the data. Hence, Greenwood et al. proposed the Greenwood–Tarazano spectrum [35], based on experimental data:

$$\Phi_n(\kappa) = 0.33C_n^2\left(\kappa^2 + \kappa\kappa_0\right)^{-11/6}, 0 \le \kappa \le \infty, \kappa_0 = 2\pi/L_0. \tag{14.21}$$

Equation (14.21) combines experimental-measurement data based on Eq. (14.18) and replaces the internal-scale parameter κ with external-scale parameter κ_0. Therefore, compared with the von Kármán spectrum, the Greenwood–Tarazano spectrum decreases faster in the inertial subrange [39].

Hill Spectrum

Strictly speaking, the spectral models represented by Eqs. (14.16, 14.17, 14.18, 14.19, 14.20, and 14.21) are correct only in the inertial subrange because their application scope in the noninertial turbulence subrange is only in the mathematical sense. They cannot be reasonably explained in the physical sense. Champagne et al. [54] proved, according to experimental measurements, that the power-spectrum density has a small protrusion feature when it is close to $1/l_0$ of the high wavenumber (see Fig. 14.15); however, the power spectrum described in Eqs. (14.18, 14.19, 14.20, and 14.21) does not have this feature.

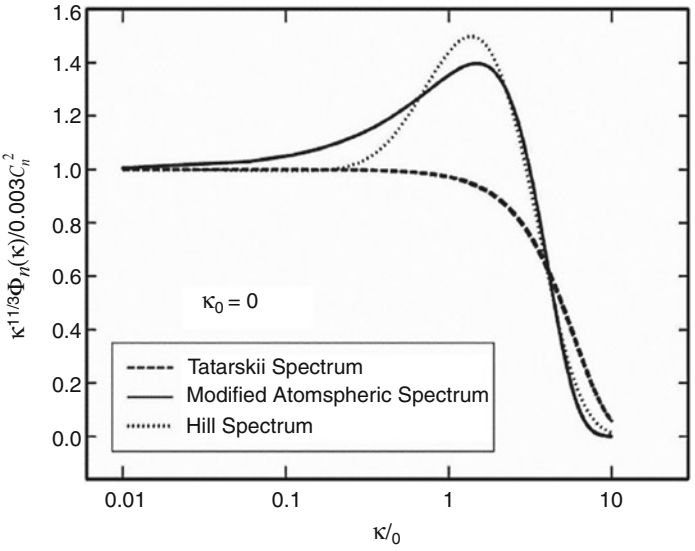


Fig. 14.15 Power-spectral density curves [63]

Figure 14.15 shows the Tatarskii, Andrews, and Hill spectral-density curves. Unlike the continuous decline of the Tatarskii spectrum, in the Hill and Andrews spectra, $\kappa = -1/l_0$; hence, the curve protrudes at $1/l_0$, and then drops rapidly.

Based on experiments, Hill considered the influence of mutation factors in the high-wavenumber region of the power spectrum; that is, the power spectrum changes from the inertial-convection range to the viscous-convection range in the relatively high-wavenumber region, and combines the κ –11/3 power law with the κ –3 power law. He then proposed the Hill spectrum, a numerical model that accurately described the inertial and dissipation regions [33]:

$$\begin{aligned}&\Phi_n(\kappa) = 0.033C_n^2\kappa^{-11/3}\\&\times\left\{\exp\left(-1.2\kappa^2 l_0^2\right) + 1.45\exp\left[-0.97(\ln\kappa l_0 - 0.452)^2\right]\right\}, 1/L_0 \le \kappa \le 1/l_0,\end{aligned} \tag{14.22}$$

where exp represents the exponential function with a constant e as the base, and ln represents the logarithm with a constant e as the base. The Hill spectrum is the solution to the following second-order linear variance:

$$\begin{aligned}&\frac{d}{dK}\left\{K^{14/3}\left[13.9K(0.135l_0)^{3.8} + 1\right]^{-0.175}\frac{d}{dK}\Phi_n(K)\right\}\\&\quad = 1.41K^4(0.135l_0)^{4/3}\Phi_n(K).\end{aligned} \tag{14.23}$$

Equation (14.23) is the second-order linear variance obtained by Hill in analyzing the experimental data, and the Hill spectrum is the numerical solution of Eq. (14.23). Because of the effect of the internal-turbulence scale, l_0, this spectrum differs from the Kolmogorov spectrum. Before entering the fast exponential decline, there is no rapid decline, but a small peak in the rise.

Andrews Spectrum

Owing to the complex form of the Hill spectral model, it is inconvenient for theoretical calculations [14]. Andrews interpolated the Hill spectrum model to provide the same theoretical model as the von Kármán turbulence spectrum. It describes the protrusion at high wavenumbers and has a mathematical form that is easy to handle, as shown in Eq. (14.24) [40]:

$$\Phi_n(\kappa) = 0.033C_n^2\left[1 + a_1(\kappa/\kappa_l) - a_2(\kappa/\kappa_l)^{7/6}\right]\frac{\exp\left(-\kappa^2/\kappa_l^2\right)}{\left(\kappa_0^2 + \kappa^2\right)^{11/6}}, \tag{14.24}$$

where $a_1 = 1.802$, $a_2 = 0.254$ and $\kappa_l = 3.3/l_0$. From Eq. (14.24), it can be seen that the Andrews spectral model can be simplified to the von Kármán spectral model

when $\alpha_1 = \alpha_2 = 0$. Moreover, the Andrews and von Kármán spectral models can be changed to the Kolmogorov spectral model when $\kappa_0 = l_0 = 0$.

The von Kármán spectrum is essentially the same as the Kolmogorov spectrum in the inertial subrange. With an increase in the wavenumber, the von Kármán spectrum decreases sharply and monotonically. The Hill and Andrews spectra differ significantly from the Kolmogorov spectrum in the high-wavenumber region of the inertial subregion, reflecting the impact of high-wavenumber mutations on optical transmissions.

Figure 14.16 shows the variation of the atmospheric refractive-index power spectrum with the spatial wavenumber. In Fig. 14.16, the outer scale $L_0 = 10$ m, inner scale $l_0 = 1$ cm, and the Kolmogorov, von Kármán, and Andrews spectra decrease with an increase in the spatial wavenumber. In the inertial-turbulence subregion, the three types of turbulence spectra exhibit the same trend, following the "−11/3" rule. When the spatial wavenumber increases to reach the dissipation region, the fluctuation of the Andrews spectrum in the dissipation region is well suppressed. However, the Kolmogorov and von Kármán spectra are unstable in the high-frequency turbulent dissipation region.

Non-Kolmogorov Spectrum

The classic Kolmogorov theory, based on the Richardson energy cascade, indicates that all turbulent vortices are isotropic in the inertial subscale range. However, in a free atmosphere, especially in a stable stratified stratosphere, turbulence may become

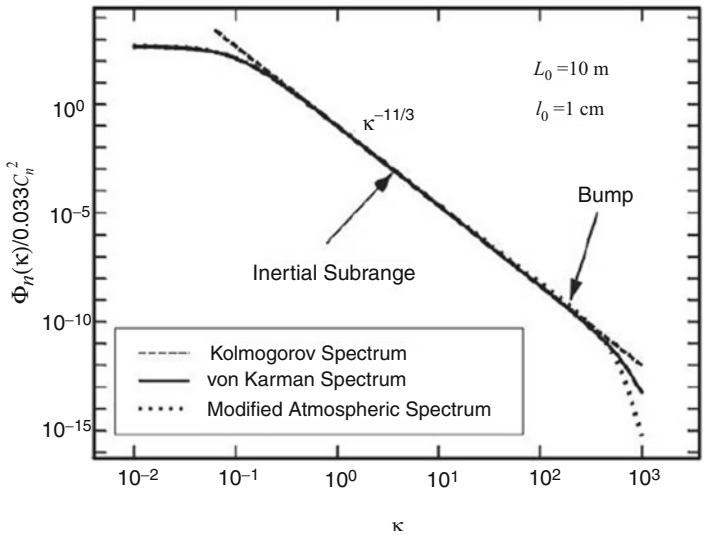


Fig. 14.16 Variation of the atmospheric refractive-index power spectrum with the spatial wavenumber [63]

anisotropic on a large scale, and most anisotropic cases occur in the boundary layer, namely, the ground, buildings, aircraft, etc.

Turbulence does not always follow the Kolmogorov power-spectrum model [5], even within the inertial range. This is due to the extraction of turbulent energy by buoyancy, resulting in a sharp decline in the energy-transfer rate. Therefore, kinetic energy is transmitted in the entire spectrum at a nonuniform rate and decreases with increasing wavenumber.

In 1962, Obukhov et al. found through experimental observations that large-scale turbulence had obvious anisotropy [76]; however, it was limited to the scope of large-scale turbulence. Later, Toselli et al. found that energy could be directly transferred between large-scale and small-scale turbulence [57]. Today, an increasing amount of experimental data shows that in actual atmospheric turbulence, most turbulence is in an anisotropic state, namely, non-Kolmogorov turbulence [88].

To study the influence of atmospheric turbulence on light-wave propagation in the non-Kolmogorov spectrum, Andrews et al. established a non-Kolmogorov spectral model [38]:

$$\Phi_n(\kappa,\alpha) = A(\alpha)\widetilde{C}_n^2\kappa^{-\infty}, 3 < \alpha < 4, \tag{14.25}$$

where κ is the spatial wavenumber, α is the power index of the power-spectral density of the refractive-index fluctuation, $\widetilde{C}_n^2$ is the refractive-index structure constant of non-Kolmogorov turbulence, in $m^{3-\alpha}$. $A(\alpha)$ maintains the consistency between the refractive-index structure function and its power spectrum: $A(\alpha) = \Gamma(\alpha - 1)\cos(\alpha\pi/2)/(4\pi^2)$.

At this time, $\alpha = 11/3$ h, $\widetilde{C}_n^2$, and $A(\alpha) = 0.033$, and Eq. (14.20) becomes the power-spectral density model of the refractive-index fluctuation of Kolmogorov turbulence.

To understand the influence of internal and external scales on the non-Kolmogorov atmospheric-turbulence effect, Andrews et al. developed a generalized modified non-Kolmogorov spectral model [32]:

$$\Phi_n(\kappa,\alpha) = \frac{A(\alpha)\widetilde{C}_n^2}{\left(\kappa^2+\kappa_0^2\right)^{\alpha/2}}\exp\left(-\kappa^2/\kappa_m^2\right), \quad 0 < \kappa < \infty, 3 < \alpha < 4, \tag{14.26}$$

where $\kappa_0 = 2\pi/L_0$, $\kappa_m = c(\alpha)/l_0$, and $c(\alpha)$ are the functions for the power of the refractive-index fluctuation power-spectral density α, which have the following forms:

$$A(\alpha) = \frac{1}{4\pi^2}\Gamma(\alpha-1)\cos\left(\frac{\pi}{2}\alpha\right) \tag{14.27}$$

$$c(\alpha) = \left[\frac{2\pi}{3}\Gamma\left(5-\frac{\alpha}{2}\right)A(\alpha)\right]^{1/(\alpha-5)}. \tag{14.28}$$

It can be seen from Eq. (14.27) that when $\alpha = 11/3$ h and $A(\alpha) = 0.033$, the non-Kolmogorov spectrum is equivalent to the modified von Kármán spectrum. When $\alpha = 11/3$, $A(\alpha) = 0.033$, $L_0 \to \infty$ and $l_0 \to 0$, the non-Kolmogorov spectrum is equivalent to the Kolmogorov spectrum.

As shown in Fig. 14.17, when $\alpha = 11/3$, $\tilde{C}_n^2 = C_n^2$, and $A(\alpha) = 0.033$, then, $\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-11/3}$ and the non-Kolmogorov spectrum becomes the Kolmogorov spectrum. When $\alpha = 3.1$, $\alpha = 10/3$ and $\alpha = 4.8$, the non-Kolmogorov spectrum obviously deviates from the Kolmogorov spectrum, and the greater the deviation, the greater the degree of deviation. However, in the inertial region (i.e., $\kappa < 1.1 \times 10^3$), the non-Kolmogorov and Kolmogorov spectra deviate slightly at different values. From the above analysis, we can see that the Kolmogorov spectrum is a special case of the non-Kolmogorov spectrum.

Figure 14.18 shows the curve of parameter M varying with the generalized exponential factor α when the beam propagates horizontally in the atmospheric-turbulent non-Kolmogorov spectrum at different propagation distances z.

Figure 14.18 shows the influence of the generalized exponential factor α on parameter M at different transmission distances. It also shows that the size of α directly affects the beam-transmission characteristics in atmospheric turbulence; that is, parameter M directly reflects the atmospheric-turbulence intensity, and the larger parameter M, the stronger the atmospheric turbulence [89]. As can be seen from

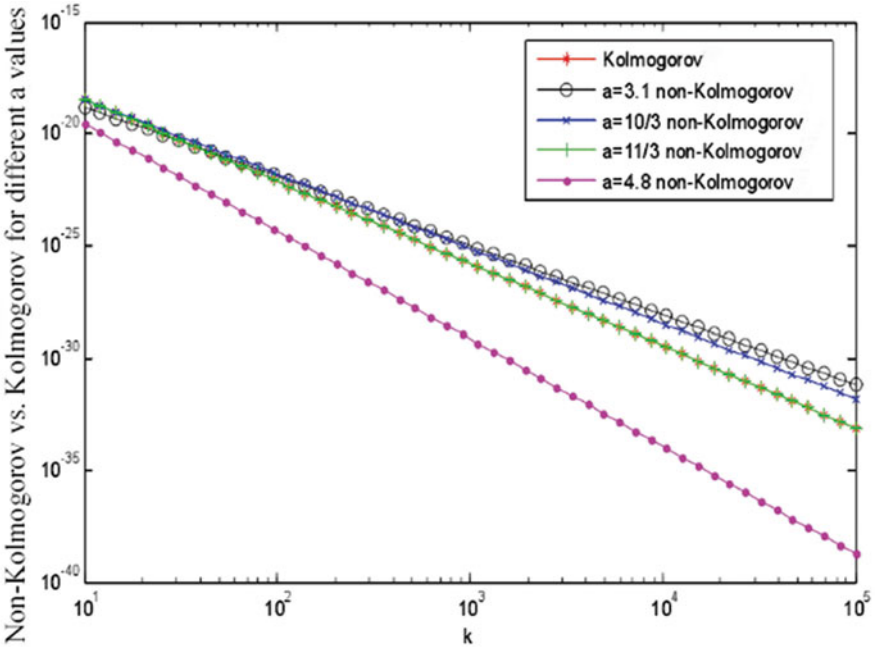


Fig. 14.17 Comparison of non-Kolmogorov and Kolmogorov spectra under different α values [89]

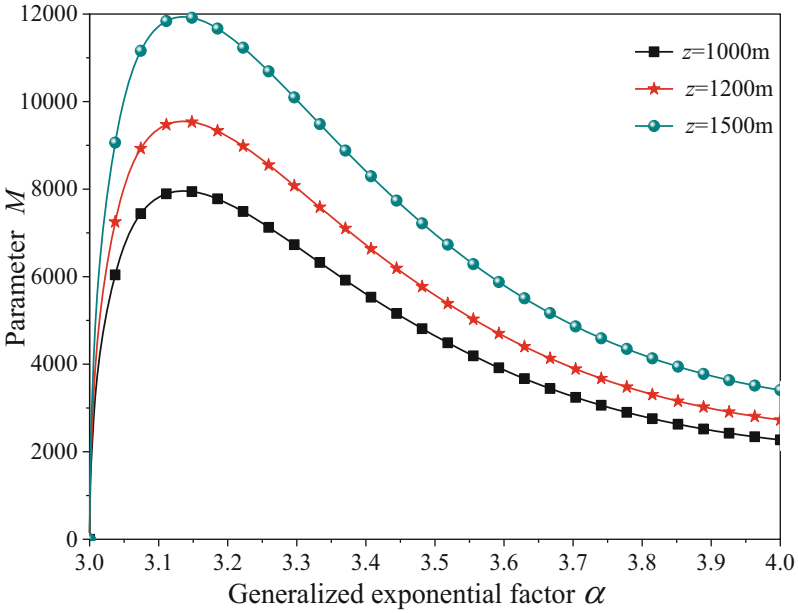


Fig. 14.18 Curve of M versus α at different transmission distances z [89]

Fig. 14.18, when z is 1000 m, 1200 m, and 1500 m, the α corresponding to the maximum M value increases correspondingly; however, it remains at approximately $\alpha = 3.1$. When $\alpha > 3.8$, parameter M tends to flatten with α.

Turbulence Spectra Varying with Altitude

There are two models of atmospheric-turbulence structure:

1. A two-layer model; that is, the turbulence in the Earth's atmosphere is composed of Kolmogorov turbulence in the troposphere and non-Kolmogorov turbulence in the stratosphere.
2. A three-layer model; that is, the turbulence in the Earth's atmosphere is composed of Kolmogorov turbulence in the boundary layer, non-Kolmogorov turbulence in the free atmosphere, and non-Kolmogorov turbulence in the stratosphere.

They have the following forms, respectively:

$$\Phi_{nB}(\kappa, z) = 0.033 C_n^2(z) \kappa^{-11/3} \tag{14.29}$$

$$\Phi_{nF}(\kappa, z) = 0.015\widetilde{C}_{nF}^{2}(z)\kappa^{-10/3} \tag{14.30}$$

$$\Phi_{nS}(\kappa, z) = 0.0024\widetilde{C}_{nS}^{2}(z)\kappa^{-5}, \tag{14.31}$$

where z is the propagation distance, $C_n^2(z)$ is the normal Kolmogorov turbulent refractive-index structural parameter in the boundary layer, with the unit of $\mathrm{m}^{-2/3}$, $\widetilde{C}_{nF}^{2}(z)$ is the non-Kolmogorov refractive-index structural parameter in the free troposphere, with the unit of $\mathrm{m}^{-1/3}$, and $\widetilde{C}_{nS}^{2}(z)$ is the non-Kolmogorov refractive-index structural parameter in the stratosphere, with the unit of m^{-2}. The above three parameters depend on the altitude.

The three-dimensional power spectrum of any power-law refractive-index fluctuation can be defined as

$$\Phi_n(\kappa, a) = A(\alpha) \cdot C_n^2 \cdot K^{-\alpha}, \tag{14.32}$$

where C_n^2 is the general refractive-index structure constant, and the unit is $m^{3-\alpha}$; $A(\alpha)$ is a constant that maintains the consistency between the structure function and its power spectrum.

According to the experiment, Arkadi found that the power spectrum of turbulence in the free troposphere and stratosphere might exhibit non-Kolmogorov characteristics [52]. The spectral behavior of the refractive-index fluctuation changes with altitude, which makes general models, such as those in Eqs. (14.21, 14.22, 14.23, 14.24, 14.25, 14.26, 14.27, 14.28, and 14.29), inapplicable.

Based on the three-layer model of the refractive-index fluctuation spectrum in the free atmosphere (boundary layer (2–3 km), free troposphere (8–10 km), and stratosphere above them), and on the experimental and theoretical data of the turbulence profile in the troposphere and lower stratosphere, Arkadi proposed a turbulence-spectrum model that varied with height [28]:

$$\alpha(z) = \frac{\alpha_1}{1 + (z/H_1)^{b_1}} + \frac{\alpha_2 (z/H_1)^{b_1}}{1 + (z/H_1)^{b_1}} \cdot \frac{1}{1 + (z/H_2)^{b_2}} + \frac{\alpha_3 (z/H_2)^{b_2}}{1 + (z/H_2)^{b_2}}, \tag{14.33}$$

where $\alpha_1 = 11/3$, $\alpha_2 = -10/3$, $\alpha_3 = -5$, z is the height from the ground, H_1 and H_2 are the characteristic heights, and b_1 and b_2 are fitting coefficients.

The experimental results showed that the turbulence characteristics do not change smoothly with height, and the height dependence of α is expressed by a three-layer height model, wherein the first layer corresponds to the boundary layer with Kolmogorov turbulence ($\alpha = 11/3$), the second corresponds to the free troposphere with non-Kolmogorov turbulence ($\alpha = -10/3$), and the third corresponds to the stratosphere with non-Kolmogorov turbulence ($\alpha = -5$).

Figure 14.19 shows the power-profile distribution of the turbulent power spectrum when $H_1 = 2$ km, $H_2 = 8$ km, $b_1 = 8$, $b_2 = 10$ (solid line), $b_1 = 15$, and $b_2 = 20$ (dotted line).

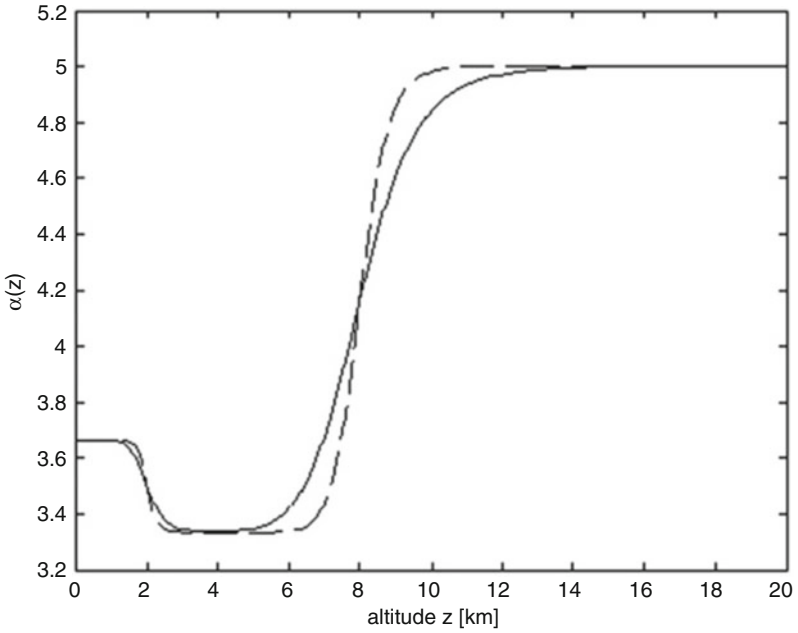


Fig. 14.19 Turbulence power-spectrum power-exponential profile distribution [52]

Atmospheric-Turbulence Channel Models

Owing to the effect of atmospheric turbulence, when a laser beam is transmitted in the atmosphere, the light intensity at the receiving end fluctuates randomly; that is, it produces the light-intensity flicker effect, which leads to signal fading at the receiver and deteriorating received-signal quality. Therefore, to improve the performance of a free-space optical (FSO) system, it is necessary to establish channel models that describe different atmospheric-turbulence conditions. Table 14.3 shows the classification of various atmospheric-turbulence channel models.

Log-Normal Distribution

The log-normal distribution probability-density function describes the flicker and fading statistics in the weakly turbulent region. Considering the log-normal model, the probability-density function of the received optical signal is given by the following formula [3]:

$$P(I) = \frac{1}{\sqrt{2\pi\sigma_l^2 I}} \exp\left\{-\frac{(\ln(I/I_0) - E[I])^2}{2\sigma_l^2}\right\}, \tag{14.34}$$

Table 14.3 Classification diagram of atmospheric-turbulence channel models

Channel-model distribution	Probability-distribution function	Applicable turbulence range
Log-normal	$P(I) = \frac{1}{\sqrt{2\pi\sigma_l^2} I} \exp\left\{ - \frac{(\ln(I/I_0) - E[I])^2}{2\sigma_l^2} \right\}$	Weak
Gamma–gamma	$p(I) = \int_0^{+\infty} p_y(I\|x) p_x(x) dx = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} I^{(\alpha+\beta)/2-1} K_{\alpha_\beta}\left[2(\alpha\beta I)^{1/2}\right]$	Medium and strong
Negative exponential	$p(I) = \frac{1}{I_0} \exp\left(-\frac{I}{I_0}\right)$	Strong
K	$P(I) = \frac{2a^{(a+1)/2}}{\Gamma(a)} I^{(a-1)/2} K_{a-1}\left(2\sqrt{aI}\right)$	Medium and strong
Malaga	$P(I) = A \sum_{k=1}^{\infty} a_k I^{\frac{\alpha+k}{2}-1} K_{\alpha-k}\left(2\sqrt{\frac{\alpha I}{\gamma\beta+\Omega}}\right)$	Weak, medium, and strong
Ke	$P(I) = A_1 \times \frac{1}{\sqrt{2\pi}\times\sigma_1} \times \mathrm{e}^{\left[-\frac{(I-\mu_1)^2}{2\times{\sigma_1}^2}\right]} + A_2 \times \frac{1}{\sqrt{2\pi}\times\sigma_2} \times \mathrm{e}^{\left[-\frac{(I-\mu_2)^2}{2\times{\sigma_2}^2}\right]}$	Weak

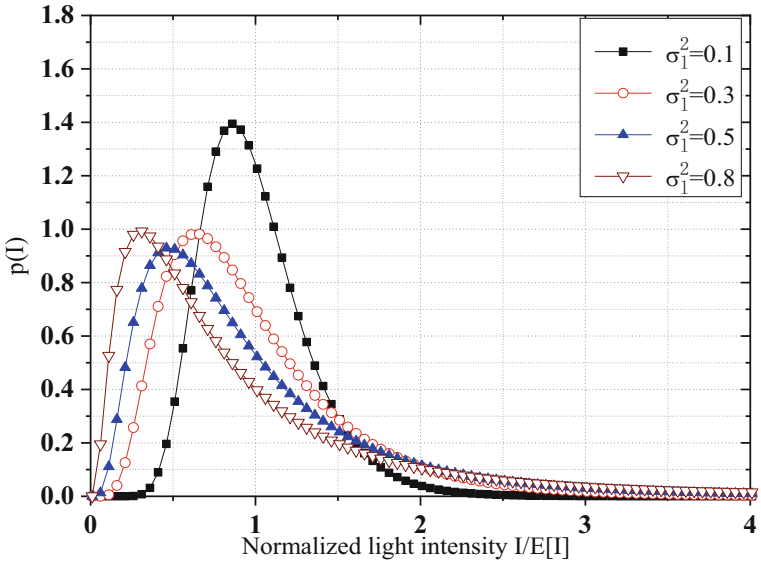


Fig. 14.20 Log-normal distribution probability-density function curve [3]

where σ_l^2 is the logarithmic light-intensity variance and I is the normalized light-intensity value reaching each receiver.

Figure 14.20 shows the probability-density function curve of the log-normal distribution under different σ_l^2 conditions. As can be seen from Fig. 14.20, with the increase in turbulence intensity σ_I^2, the curve gradually deviates from the mean value, and the tail becomes more serious, which is gradually inconsistent with the log-normal distribution. The log-normal distribution is only applicable to weak-turbulence conditions ($\sigma_I^2 < 1$) because the scattering effect is also considered with the increase in turbulence intensity.

Gamma–Gamma Distribution

Because the log-normal model is limited to weak turbulence, Shah produced another model based on the modified Rytov theory, called the gamma–gamma channel model [89], in which the light-intensity fluctuation was developed as the product of two random processes and the gamma–gamma distribution:

$$p(I) = \int_0^{+\infty} p_y(I|x)p_x(x)dx = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} I^{(\alpha+\beta)/2-1} K_{\alpha-\beta}\left[2(\alpha\beta I)^{1/2}\right], \quad I > 0\,. \quad (14.35)$$

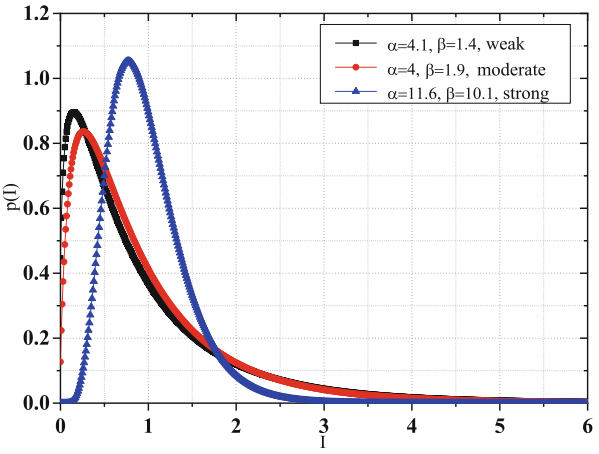


Fig. 14.21 Gamma–gamma probability-density function of distributed light intensity [3]

In the formula, the parameters α and β represent the large- and small-scale scattering coefficients, respectively, $K_n(\cdot)$ represents the second type of correction function, and $\Gamma(\cdot)$ represents the gamma function.

Figure 14.21 shows the probability-density function of the gamma–gamma distribution under different turbulence intensities. It can be seen from the figure that, compared with the log-normal distribution, the gamma–gamma distribution has a wider application scope and can accurately describe the statistical characteristics of light-intensity fluctuations in weak, medium, and strong turbulence regions.

As shown in Fig. 14.22, the scattering coefficients of the large and small scales are large when σ_I^2 is small. When σ_I^2 gradually increases to approximately 3, the scattering coefficients of the large and small scales decrease rapidly. When σ_I^2 gradually increases from 1 to 100, α increases, and β decreases and approaches 1.

Negative-Exponential Distribution

In the case of strong turbulence with a link length of several kilometers, an increase in the scattering number leads to fading, which is considered a random process. In this case, the signal amplitude follows the negative-exponential distribution of the signal strength. The negative-exponential probability-distribution function is

$$p(I) = \frac{1}{I_0} \exp\left(-\frac{I}{I_0}\right), \qquad I > 0, \tag{14.36}$$

where $I_0 = E[I]$ denotes the average light intensity. In the saturated state, the flicker index→1, and the negative-exponential distribution is shown in Fig. 14.23.

Figure 14.24 shows the light-intensity fluctuation probability-distribution curve of the gamma–gamma and log-normal distributions with different Rytov variances.

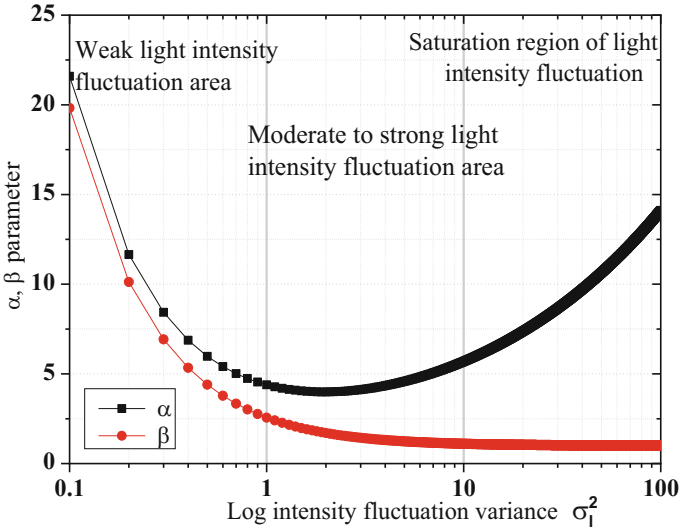


Fig. 14.22 Summation of α and β values under different σ_I^2 conditions [3]

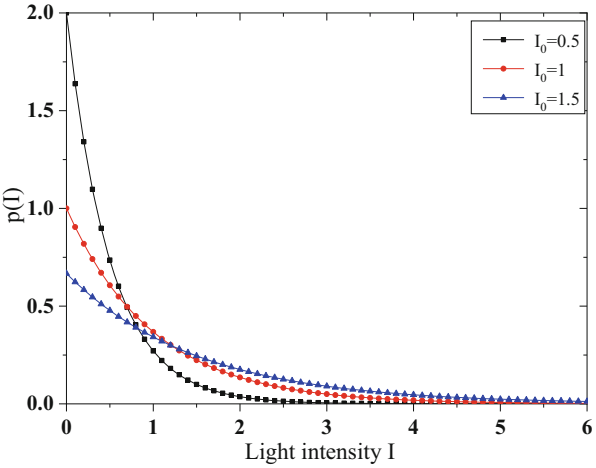


Fig. 14.23 Probability-density function of the negative-exponential distribution light-intensity fluctuation [89]

In the weak turbulent region, the log-normal distribution model is in good agreement with the experimental data; however, in the strong turbulent region, the tail phenomenon of the log-normal distribution model is significant, which deviates greatly from the experimental data. In the weak, medium, and strong turbulent regions, the gamma–gamma distribution model is more consistent.

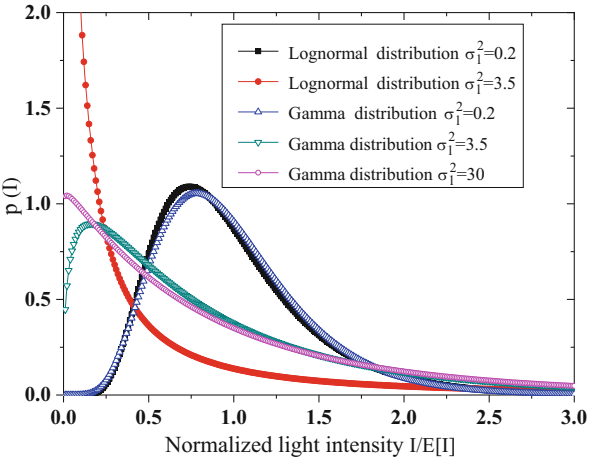


Fig. 14.24 Probability-density function of light-intensity fluctuation [89]

K Distribution

In the strong turbulence region, the atmospheric-turbulence channel model can be represented by the K distribution, and its probability-density function is [89]

$$f_I(I) = \frac{2a^{(a+1)/2}}{\Gamma(a)} I^{(a-1)/2} K_{a-1}\left(2\sqrt{aI}\right), \quad I > 0. \tag{14.37}$$

The K distribution function can be expressed as

$$F(S,a) = \int_{-\infty}^{s} \frac{2a^{(a+1)/2}}{\Gamma(a)} S^{(a-1)/2} K_{a-1}\left(2\sqrt{as}\right) dS, \tag{14.38}$$

where S is a random variable.

Figure 14.25a shows the slot error rate of the K-distribution channel with the SNR. It can be observed from Fig. 14.25a that when the SNR is small, the slot error rate of the two-dimensional K-distribution channel is lower than that of the one-dimensional K-distribution channel, whereas when the SNR is large, the difference between the two is not significant.

Figure 14.25b shows the channel capacity of the K-distribution channel changing with SNR when $a = 1$ and $a = 2$. It can be seen from Fig. 14.25b that the increase in SNR and a will increase the average channel capacity of the K distribution. When a is the same, the two-dimensional K-distribution channel has a larger capacity than the one-dimensional K-distribution channel [90].

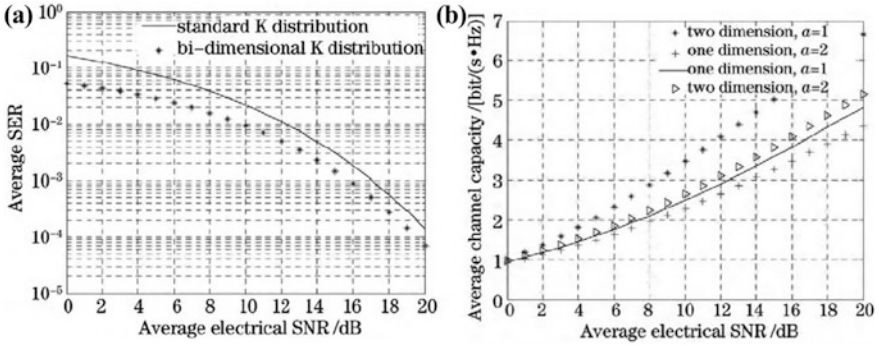


Fig. 14.25 K-distribution channel characteristics. (**a**) Slot error rate of K-distribution channel with SNR [91]. (**b**) Channel capacity of the K distribution channel varies with SNR when $a = 1$ and $a = 2$ [91]

Malaga Distribution

The Malaga channel model can describe various turbulence distributions, and its probability-density function is [91]

$$f_I(I) = A\sum_{k=1}^{\infty} a_k I^{\frac{\alpha+k}{2}-1} K_{\alpha-k}\left(2\sqrt{\frac{\alpha I}{\gamma\beta+\Omega'}}\right) \tag{14.39}$$

$$\begin{cases} A = \dfrac{2\alpha^{\frac{\alpha}{2}}}{\gamma^{1+\frac{\alpha}{2}}\Gamma(\alpha)}\left(\dfrac{\gamma\beta}{\gamma\beta+\Omega'}\right)^{\beta} \\ a_k = \dfrac{(\beta)_{k-1}(\alpha\gamma)^{\frac{k}{2}}}{[(k-1)!]^2\gamma^{k-1}(\Omega'+\gamma\beta)^{k-1}} \end{cases}, \tag{14.40}$$

where a and β are expressed as positive real and fading parameters, respectively, and the Ω' parameter is the coherent average optical power, φ_A represents the phase of the light-wave line-of-sight propagation component, and φ_B represents the phase of the light-wave line-of-sight propagation coupling component [92].

It can be seen from Fig. 14.26(a) that as the value of ρ increases, the value of $f_I(I)$ in the ellipse box correspondingly increases. In the case of strong turbulence, compared with the gamma–gamma distribution, $f_I(I)$ of the K distribution has a relatively slow downward trend and a relatively serious tail. It can be seen from Fig. 14.26(b) that in the case of weak turbulence, the $f_I(I)$ curve gradually moves to the right with an increase in the value of ρ, while the peak value increases and the tailing phenomenon decreases [92–94].

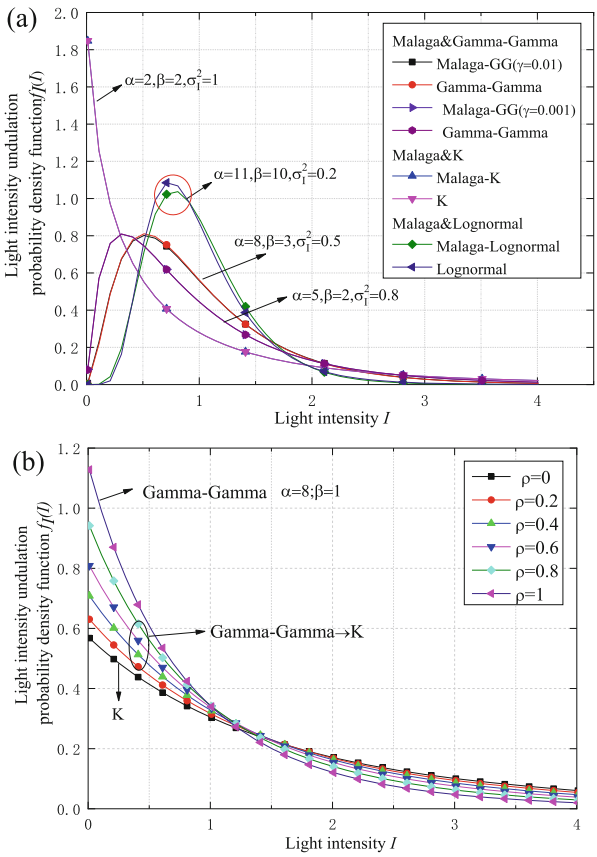


Fig. 14.26 Curve of the probability-density function of the light-intensity fluctuation of the Malaga distribution with values of ρ under different α and β conditions [92]. (**a**) $\alpha = 8$, $\beta = 1$. (**b**) $\alpha = 17$, $\beta = 16$

Ke's Model

For visible-light communication in the Internet of Vehicles, Ke Xizheng proposed a nighttime background-light noise model (Ke's model) that uses a double Gaussian function. It is called the Coriolis model [95]:

$$f(I) = A_1 \times \frac{1}{\sqrt{2\pi} \times \sigma_1} \times \mathrm{e}^{\left[-\frac{(I-\mu_1)^2}{2\times\sigma_1{}^2}\right]} + A_2 \times \frac{1}{\sqrt{2\pi} \times \sigma_2} \times \mathrm{e}^{\left[-\frac{(I-\mu_2)^2}{2\times\sigma_2{}^2}\right]} \tag{14.41}$$

Figure 14.27 shows the Coriolis distribution model. The two Gaussian curves in the figure correspond to the two Gaussian functions of Eq. (14.41), and the corresponding mean and variance are different. In Gaussian curve 1, the background noise mainly contains the current street-lamp background light component, whose

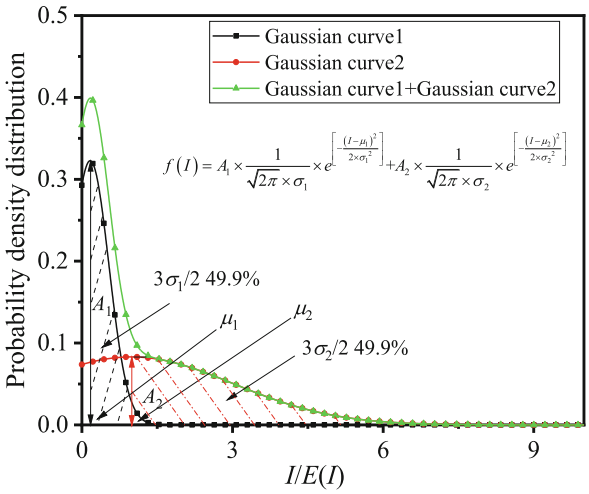


Fig. 14.27 Coriolis distribution model [95]

light intensity is weak and the light-intensity fluctuation variance is small. In Gaussian curve 2, the background noise mainly includes front and rear vehicle lights, extremely bright billboards, traffic lights, and other components with a strong intensity and large intensity-fluctuation variance [95].

Figure 14.28 shows the survey data for different cities. According to Fig. 14.28, the following can be seen:

1. The Coriolis distribution model can accurately describe the environmental noise model of visible-light communication.
2. The probability distribution for rainy days is no different from that of sunny days.
3. The light-intensity fluctuation generated by current street lamps is significantly smaller than that generated by other light sources.

Summary

It has been 81 years since the first publication of Kolmogorov's theory in 1941. During this period, scientific researchers from various countries have successively established various theoretical models of the atmospheric-turbulence spectrum. Research on the refractive-index power spectrum of atmospheric optical turbulence has developed, from the Kolmogorov refractive-index power spectrum (describing statistically uniform and fully developed ideal turbulence) to the non-Kolmogorov refractive-index power spectrum (describing nonclassic turbulence, such as stratification and turbulence that are not fully developed) [87]. However, considering the need for theoretical analysis, it is still necessary to study new reasonable turbulence-spectral models that can simplify the theoretical analysis of turbulence effects.

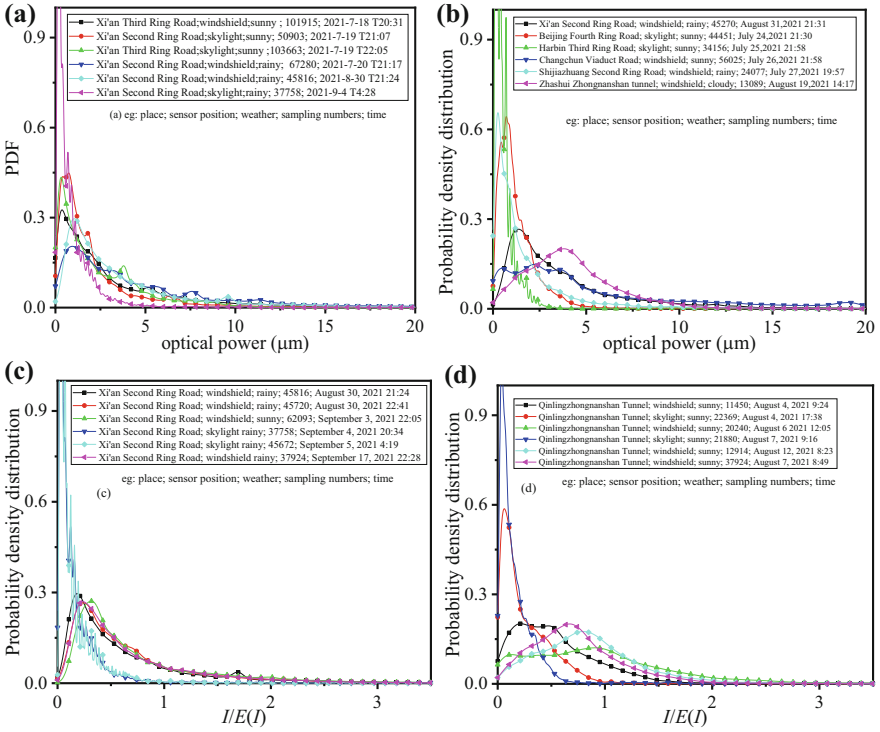


Fig. 14.28 Probability distribution of nighttime background-light noise for visible-light communication under different conditions [95]. (**a**) Different regions in Xi'an. (**b**) Different cities. (**c**) Rain. (**d**) Tunnel

This chapter systematically summarized the application and development of atmospheric-turbulence models. Researchers have been investigating and developing atmospheric turbulence in the following aspects.

1. Measuring and studying the characteristics of actual atmospheric turbulence, obtaining more statistical-characteristic parameters of atmospheric turbulence, and accelerating the research and development of optical wireless communication in actual atmospheric turbulence.
2. Analyzing and comparing the transmission characteristics of different light sources in different atmospheric-turbulence conditions, exploring the influence mechanism of atmospheric turbulence on optical signal transmission, and improving the ability to suppress atmospheric turbulence.
3. Researching high-accuracy atmospheric-turbulence parameter measurement instruments and developing high-sensitivity and high-resolution atmospheric-turbulence characteristic measurement technologies and methods to improve the accuracy of atmospheric-turbulence parameter measurements in the actual atmosphere.

References

1. Kim, John, Moin, et al. Turbulence statistics in fully developed channel flow at low Reynolds number [J]. J Fluid Mech, 1987, 50(24): 485–499.
2. A S Monin. Statistical Fluid Mechanics: The Mechanics of Turbulence [J]. American Journal of Physics, 1977, 318(10):202–220.
3. Xingzheng Ke, Ke Dong, Optical Wrieless Communication [M],Beijing:Springer Press, 2022,:10–14.
4. R Rosa. Turbulence theories [J]. Antennas & Propagation IEEE Transactions on, 2006, 45(42): 295–303.
5. D Wilson. Acoustic scattering and the spectrum of atmospheric turbulence [J]. The Journal of the Acoustical Society of America, 1999, 105(1):30–34.
6. K.S.Shaik. Atmospheric propagation effects relevant to optical communication [R]. TDA Progress Report, 1989.
7. Wyngaard J C, Izumi Y, Collins S A, et al. Behavior of the refractive-index-structure parameter near the ground [J]. Journal of the Optical Society America A, 1971, 61(12):1646–1650.
8. Neff W D. Quantitative evaluation of acoustic echoes from the planetary boundary layer [R]. PCAI Progress Report, 1975.
9. Kukharets V.P, Tsvang L.R. Atmospheric Turbulence Characteristics Over a Temperature-Inhomogeneous Land Surface. Part I: Statistical Characteristics of Small-Scale Spatial Inhomogeneities of Land Surface Temperature [J]. Boundary-Layer Meteorology, 1998, 86(1): 89–101.
10. Li H S. Effect of turbulence on laser beam transmission in the atmosphere [D]. Changchun: Changchun University of Technology, 2002.
11. Hufnagel R E. Variations of Atmospheric Turbulence [C]//Topical Meeting on Optical Propagation Through Turbulence in Proe IEEE. Boulder, 1974:76–80.
12. Hufnagel R E, Stanley N R. Modulation Transfer Function Associated with Image Transmission through Turbulent Media [J]. Journal of the Optical Society of America, 1964, 54(1):52–60.
13. R.L. Fante. Electromagnetic Beam Propagation in Turbulent Media: An Update [J]. Proceeding of the IEEE, 1980, 68(11): 1424–1443
14. ITU-R. Propagation data and prediction methods required for the design of earth-space telecommunication systems [C]//Radio Communication Stuty Group Meeting, Budapest, 2001: 618–620.
15. Linyan Cui, Wen Fei, and Bindang Xue. Theoretical investigations of infrared optical wave modulation transfer function models in anisotropic marine turbulence [J]. Journal of the Optical Society of America A, 2018, 35(9):1496–1503.
16. Gao C, Yu S Y, Ma J, et al. Beam wander in moderate to strong turbulence [J]. High Power Laser and Particle beam, 2006, 18(10):1597–1601.
17. Yang R K, Liu K X, Wu Z S, etc. Study on the influence of the outer scale of turbulent atmosphere on the scintillation of beam-ground propagation [J] Journal of Radio Wave Science, 2006, 21(6):91–95.
18. Kunkel K E, Walters D L. Modeling the diurnal dependence of the optical refractive index structure parameter [J]. Journal of Geophysical Research Oceans, 1983, 88(C15):10999–11004.
19. Fante, R.L., Some physical insights into beam propagation in strong turbulence [J]. Radio Science, 1980, 15(4):757–762.
20. Fante, R.L., Two-position, two-frequency mutual-coherence function in turbulence [J]. Journal of the Optical Society of America, 1981, 71(12):1446–1451.
21. Fante, R.L., Inner-scale size effect on the scintillations of light in the turbulent atmosphere [J]. Journal of the Optical Society of America, 1983, 73(3):277–281.
22. Cheng C F, Qi D P, Liu D L, etc Simulated generation of Gaussian correlation random surface and its light scattering speckle field and probability analysis of light intensity [J] Journal of Physics, 1999, 48(09): 1635–1643.

23. Chen X Y, Teng S Y, Song H S, etc. Speckle intensity probability density in Fresnel deep region [J] Journal of Shandong Normal University: Natural Science Edition, 2006, 21(2): 48–410.
24. Vetelino F S, Young C, Andrews L, et al. Aperture averaging effects on the probability density of irradiance fluctuations in moderate-to-strong turbulence [J]. Applied Optics, 2007, 46(11): 2099–2109.
25. Tiker A, Yarkoni N, Blaunstein N, et al. Prediction of data stream parameters in atmospheric turbulent wireless communication links [J]. Applied Optics, 2007, 46(2):190–1910.
26. Zhang Y X, Du Y J. Spectral reflection and transmission model of halftone fluorescence image [J]. Journal of Optics, 2007, 27(2):365–370.
27. Hu F. Turbulence Intermittence and Atmospheric Boundary Layer [M]. Beijing: Science Press, 1995, 89–92.
28. Chemov L A. Wave Propagation in A Random Medium [M]. New York: McGraw-Hill, 1960, 56–59.
29. Zhang Y X, Chi Z Y. Transmission and imaging of light waves in the atmosphere [M] Beijing: National Defense Industry Press, 1997, 17–19
30. A N Kolmogorov. The local structure of turbulence in an incompressible viscous fluid for very large Reynolds numbers [J]. Soviet Physics Uspekhi, 1941, 30(42): 301–305.
31. V T Karman. Progress in the Statistical Theory of Turbulence [J]. Proceedings of the National Academy of ences of the United States of America, 1948, 34(11):530–5310.
32. V I Tatarski. Wave Propagation in a Turbulent Medium [M]. New York: Hill Pergamon Press, 1961: 69–79.
33. V T Karman. From low-speed aerodynamics to astronautics [M]. New York: Pergamon Press, 1963: 152–169.
34. N E Busch, H A Panofsky. Recent spectra of atmospheric turbulence [J]. Quarterly Journal of the Royal Meteorological Society, 1968, 94(400): 132–149.
35. J C Kaimal, J C Wyngaard, Y Izumi, et al. Spectral characteristics of surface layer turbulence [J]. Quarterly Journal of the Royal Meteorological Society, 1972, 98(28): 12–20.
36. D P Greenwood, D Tarazano. A ProPosed form for the atmospheric microtem Perature in the input range [J]. USAF Rome Air Development Center. Technical Report RADC, 1974, 42(16): 442–456.
37. F H Champagne. Flux Measurements, Flux Estimation Techniques, and Fine-Scale Turbulence Measurements in the Unstable Surface Layer Over Land [J]. J. Opt. Soc. Am, 1977, 34(3): 515–530.
38. R J Hill. Modified spectrum of atmospheric temperature fluctuations and its application to optical propagation [J]. J. Opt. Soc. Am, 1978, 68(04): 892–897
39. R M Manning. An anisotropic turbulence model for wave propagation near the surface of the Earth [J]. Antennas & Propagation IEEE Transactions on, 1986, 34(2):258–261.
40. R H Kraichnan. Preliminary Calculation of the Kolmogorov Turbulence Spectrum [J]. The Physics of Fluids, 1990, 8(5): 995–998.
41. L C Andrews. An analytical model for the refractive index power spectrum and its application to optical scintillations in the atmosphere [J]. Journal of Modern Optics, 1992, 39(9): 1849–1853.
42. R R Beland. Some aspects of propagation through weak isotropic non-Kolmogorov turbulence [J]. Proceedings of SPIE - The International Society for Optical Engineering, 1995, 27 (11): 6–16.
43. L V Antoshkin, N N Botygina, O N Emaleev, et al. Investigation of turbulence spectrum anisotropy in the ground atmospheric layer: Preliminary results. 1995, 53(2): 54–62.
44. B E Stribling, B M Welsh, M C Roggemann. Optical Propagation in Non-Kolmogorov Atmospheric Turbulence [J]. Proceedings of SPIE The International Society for Optical Engineering, 1998, 24 (7):117–126.
45. M S Belen'Kii, S J Karis, C L Osmon, et al. Experimental evidence of the effects of non-Kolmogorov turbulence and anisotropy of turbulence [J]. Proceedings of SPIE - The International Society for Optical Engineering, 1999, 44 (7):37–46.

46. C Lipatack. Optical wave propagation through non-Kolmogorov atmospheric turbulence [M]. University of Central Florida, 2004: 152–1510.
47. I Toselli, L C Andrews, R L Phillips, et al. Scintillation index of optical plane wave propagating through non-Kolmogorov moderate strong turbulence [J]. Proceedings of SPIE The International Society for Optical Engineering, 2007, 54 (32):437–446.
48. A Zilberman, E Golbraikh, N S Kopeika. Propagation of electromagnetic waves in Kolmogorov and non-Kolmogorov atmospheric turbulence: three-layer altitude model [J]. Applied Optics, 2008, 47(34): 6385–6391.
49. I Toselli, L C Andrews, R L Phillips, et al. Free space optical system performance for laser beam propagation through non Kolmogorov turbulence for uplink and downlink paths [J]. Optical Engineering, 2008, 47(2):115–122.
50. A Zilberman, E Golbraikh, N S Kopeika. Some limitations on optical communication reliability through Kolmogorov and non-Kolmogorov turbulence [J]. Optics Communications, 2010, 283(7):1229–1235.
51. I Toselli. Introducing the concept of anisotropy at different scales for modeling optical turbulence [J]. Journal of the Optical Society of America A, 2014, 31(8): 1868–1875.
52. T Helin, J A Kinder, et al. Atmospheric turbulence profiling with unknown power spectral density [J]. Inverse Problems An International Journal of Inverse Problems Inverse Methods, 2018, 8(10): 29–43.
53. F Santos, L B Bolívar, C H Venner, et al. Modelling the Dissipation Range of von Karman Turbulence Spectrum [J]. Optics Communications, 2021, 28(11): 101–114..
54. I Goldman. Analytic derivation of the inertial range of compressible turbulence [J]. Optics Communications, 2021, 24(12): 129–140.
55. K Iwano, J Hosoi, Y Sakai, et al. Power spectrum of high Schmidt number scalar in a turbulent jet at a moderate Reynolds number [J]. Experiments in Fluids, 2021, 62(6): 148–156.
56. Zhou P Y, Huang Y N. Statistical theory of vortex structure in homogeneous isotropic turbulence [J]. Chinese science, 1975, 20(02): 180–199.
57. Liu Z X. Turbulence spectrum of atmospheric turbulence and related air pollution problems [J] Atmospheric science, 1978, 11(03):219–2210.
58. An Y Q. Energy spectral density of homogeneous isotropic turbulence in the region of large wave number [J] Journal of Hebei Institute of Technology, 1984, 22(03): 70–81.
59. Song Z F, Zhang Y X, Gong Z B. Discussion on some problems of beam drift in turbulent atmosphere [J] Journal of Quantum Electronics, 1986, 10(01): 57–64.
60. Wu Z S. Intensity correlation and scintillation of wave propagation saturation in random media [J] Journal of Radio Wave Science, 1986, 35(01): 72–84.
61. Wang J M. Spectral characteristics of near-surface atmospheric turbulence in valley cities [J] Atmospheric science, 1992, 21(01): 11–18.
62. Xu Y, Zhou C, Li Z, et al. Microstructure and Spectral Characteristics of Turbulence in the Surface- Layer Atmosphere over Guangzhou [J]. Chinese Journal of Atmospheric Sciences, 1993, 55(11): 26–210.
63. Hu F. Wavelet decomposition of turbulent vortex structure in atmospheric boundary layer [J] Climate and environmental research, 1998, 20(02):2–11.
64. Zeng Z Y, Yuan R M, Wu X Q, etc. Near surface temperature spectrum of complex terrain [J] Journal of Quantum Electronics, 1998, 32(02): 134–1310.
65. Rao R Z. Turbulence spectrum and intermittence in the study of atmospheric light propagation [J] Journal of University of Electronic Science and Technology, 1999, 28(04): 437–442.
66. Rao C H, Wen H. Spatial and temporal characterization of phase fluctuations in non-Kolmogorov atmospheric turbulence [J]. Journal of Modern Optics, 2000, 47(6): 1111–1126.
67. Rao C H, Jiang W H, Ling N. Measurement of atmospheric turbulence parameters using Hartmann-Shack wavefront sensor [J] Journal of Optics, 2000, 35(09): 1201–1208.

68. Zhu W Y, Zhao Z L, Ma X S, et al. Optical method for simultaneous measurement of refractive index structure constant and internal scale of atmospheric turbulence [J] Intense laser and particle beam, 2005, 54(10):23–26.
69. Zhang Y X. Research on wavefront distortion of laser propagation in turbulent atmosphere [D] Nanjing: Nanjing University of Technology, 2005: 32–39.
70. Liu S H, Li J, Liu H P, etc Turbulence spectrum and local isotropic characteristics in EBEX-2000 experimental data [J]. Atmospheric Science, 2005, 38(02):213–224.
71. Pan F, Ma J, Tan L Y, etc Theoretical and experimental study on atmospheric scintillation spectrum under aperture reception [J] Strong laser and particle beam, 2006, 21(09): 1457–14510.
72. Gong J D. Analysis and research on the change of two-dimensional feature length with model resolution in data assimilation [J] Atmospheric science, 2007, 20(03): 459–468.
73. He W G, Wu J, Yang C P, et al. Estimation and measurement of near-surface optical turbulence [J] Journal of University of Electronic Science and Technology, 2008, 37(06): 893–896.
74. Lu P P, Wu X Q. Study on the characteristics of turbulence spectrum on snow [J] Journal of Atmospheric and Environmental Optics, 2008, 32(06): 421–426.
75. Cheng X H, Sun F, Zhao Z W. Study on the fluctuation characteristics of atmospheric refractive index measured at sea [J] Journal of Radio Wave Science, 2010, 25(03): 513–519.
76. Ma J, Tan L, Fu S, et al. Effect of non-kolmogorov turbulence on angle-of-arrival fluctuations of starlight [J]. Journal of Russian Laser Research, 2010, 31(4): 332–338.
77. Chen H, Jin Y Q. Retrieval of atmospheric temperature and water vapor profile of FY-3 MWTS/MWHS - a case study of Typhoon Phoenix in 2008 [J] Journal of Remote Sensing, 2011, 15(01): 137–148.
78. Xue B, Cui L, Xue W, et al. Generalized modified atmospheric spectral model for optical wave propagating through non-Kolmogorov turbulence [J]. Journal of the Optical Society of America A, 2011, 28(5):912–916.
79. Li K, Pan W. Atmospheric Noises in SLF/ELF Ranges [J]. In Propagation of Electromagnetic Waves, 2014, 8(12): 254–262.
80. Duan M Y, Ai Y, Xiong Z, et al. Indoor simulation study of laser atmospheric transmission [J] Journal of Atmospheric and Environmental Optics, 2014, 9(03): 188–193.
81. Li J, Yang S, Guo L, et al. Anisotropic power spectrum of refractive-index fluctuation in hypersonic turbulence [J]. Applied Optics, 2016, 18(9):53–62.
82. Wang H Q, Li Y, Hu Q, et al. Experiment on the characteristics of light intensity fluctuation at night in Lanzhou [J] Journal of Photonics, 2018, 47(04): 194–201.
83. Zhang N N, Shan X, Zhang Y G, etc. Simulation experiment of the influence of Non-Kolmogorov turbulence on light intensity [J] Journal of Photonics, 2018, 47(6): 9–13.
84. Hu C Q, Zhang X T, Wang B Q, et al. Research on wind parameters for the design of Wuhan Qingshan Yangtze River Highway Bridge [J] Rainstorm disaster, 2019, 38(03): 276–283.
85. Zhai C. Anisotropic Non-Kolmogorov Turbulence Spectrum with Anisotropic Tilt Angle [J]. Photonics, 2021, 8(21):521–530.
86. Yang B, Yuan Q, Tan C H, et al Study on the characteristics of atmospheric turbulence exchange in the lower reaches of the Lhasa River in the eastern Qinghai-Tibet Plateau [J] Plateau meteorology, 2022, 41(01): 204–218.
87. Li B Y, Feng Y S, Teng Y J, et al. Scintillation suppression of 13.7 km horizontal channel space laser communication [J]. Optical Precision Engineering, 2022, 30(05):545–554.
88. S Karp, R M Gagliardi, S E Moran, et al. Optical Channels [M]. New York: Springer US, 1988: 105–113.
89. Rao R Z, Li Y J. Light propagation in non-Kolmogorov atmospheric turbulence and its influence on optoelectronic engineering [J] Journal of Optics, 2015, 35(05):26–36.
90. Ke X Z. Optical Wireledd Comunication [M].Springer Press, 2022: 52–59.
91. Ke X Z, Tian X C. Modeling and simulation of two-dimensional correlated K-distribution turbulent channel [J] Journal of Optics, 2015, 35(04): 41–410.

92. Wang C H. Research on phase noise characteristics and compensation technology of wireless optical subcarrier modulation [D] Xi'an: Xi'an University of Technology, 2019: 34–42.
93. Ke X Z, Wang C H, Chen Dan Phase noise analysis of subcarrier modulation system in Malaga atmospheric turbulence channel [J]. Journal of Communications, 2018, 39(11): 80–86.
94. Wang S. Optical beam propagation for a partially coherent source in the turbulent atmosphere [J]. JOSA, 1979, 69(9): 319–322.
95. Ke X Z, Qin H H, Yang S J, etc Night background light noise model of the visible light communication system of the Internet of Vehicles [J] Journal of Radio Wave Science, 2021, 36(06):986–990.

Atmospheric-Turbulence Suppression Methods in Optical Wireless Communication

15

Contents

Introduction 488
Optical Wireless Communication Systems 488
Atmospheric Turbulence 488
Atmospheric-Turbulence Formation Mechanism 488
Impact of Atmospheric Turbulence on Optical-Communication Systems 490
Key Technologies for Suppressing Atmospheric Turbulence 492
Large-Aperture Receiving Technology 492
Partially Coherent Optical Technology 495
Diversity Technology 498
Adaptive-Optics Technology 502
Development Trends 508
Summary and Outlook 509
References 510

Abstract

In optical wireless communication systems, the optical signal is affected by atmospheric turbulence during atmospheric-channel transmissions. Therefore, it is important to study methods for suppressing atmospheric turbulence in the development of optical wireless communication. This chapter analyzes the effects of atmospheric turbulence on optical wireless communication systems. It presents the current domestic and foreign research on suppressing atmospheric turbulence and finally presents future development trends and outlooks.

Keywords

Optical wireless communication · Atmospheric turbulence · Turbulence suppression

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_15

Introduction

Optical wireless communication [1] has the advantages of a high information-transmission rate, no need for frequency applications, and strong confidentiality. It has a wide range of application prospects in the future development of wireless communication. However, optical wireless communication signals are affected by atmospheric turbulence during transmission in the atmospheric channel, which produces effects, such as light-intensity fluctuation, beam expansion, beam drift, and phase fluctuation. This reduces the performance of optical wireless communication systems, making it difficult to utilize their advantages. Therefore, it is important to study the effects of atmospheric turbulence on optical communications [2].

Optical Wireless Communication Systems

A model diagram of an optical wireless communication system is shown in Fig. 15.1. As can be seen from the figure, the transmitter side is mainly used for channel coding, signal modulation, and output by a laser or amplifier and is then transmitted by a single or multiple antennas through different types of channels. The receiving end uses a single or multiple antennas for reception. After being detected, the signal passes through the spatial optical-coupling module and preamplifier. Finally, it is demodulated and decoded to produce the originally transmitted signal.

In the future, optical wireless communication will have broad market potential and unique attractions; however, optical wireless communication systems still have many areas that need improvement.

One of the main issues is the impact of the atmospheric channel on optical wireless communication. The atmospheric channel is very complex, and optical signals are affected by atmospheric turbulence during transmission. Atmospheric turbulence can lead to uncertain changes in the amplitude and phase of the received signal, which affects the stable performance of optical communication and greatly reduces the communication quality. To exploit the advantages of optical wireless communication, it is important to study methods for suppressing atmospheric turbulence.

Atmospheric Turbulence

Atmospheric-Turbulence Formation Mechanism

Factors in the atmosphere are constantly changing, such as random changes in wind speed and air pressure, which makes the atmospheric refractive index inhomogeneous, and also forms turbulent vortices [3]. Atmospheric turbulence originates from the continuous formation or dissipation of these turbulent vortices with the change in wind speed and their continuous overlapping. This produces an irregular turbulent motion.

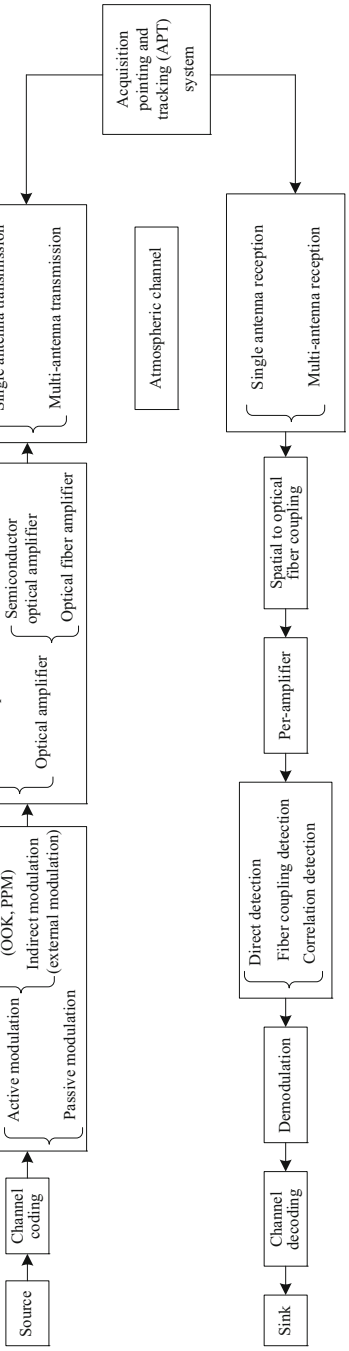


Fig. 15.1 Model of an optical wireless communication system [1]

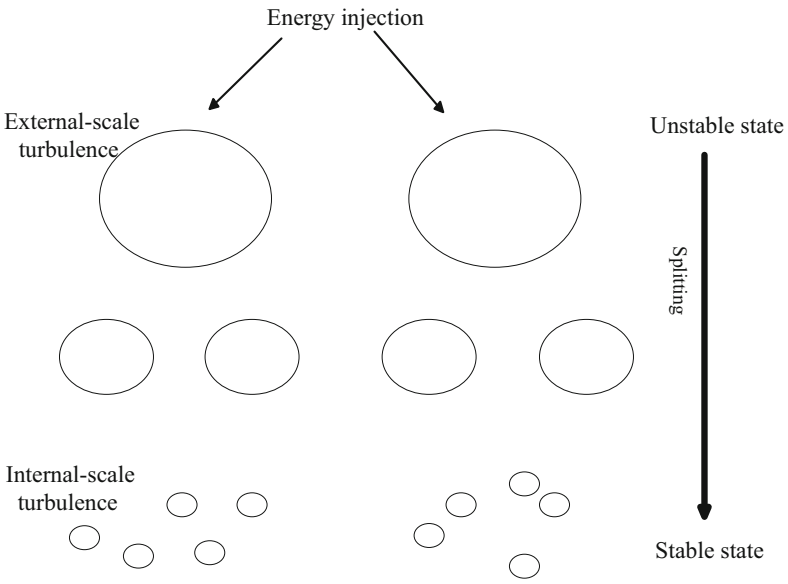


Fig. 15.2 Schematic diagram of the turbulent motion of the atmosphere [5]

The Reynolds number [4] well describes the characteristics of the turbulent energy conversion, as shown in Fig. 15.2. The outer turbulence scale has a large energy, which constitutes the initial turbulence energy. When the Reynolds number is large, it forms the outer turbulence scale. This state is very unstable so it splits into smaller turbulent vortices. The energy is dispersed to these small-scale turbulence vortices, and the splitting continues, with the Reynolds number decreasing accordingly, until it converges to a stable state.

Impact of Atmospheric Turbulence on Optical-Communication Systems

When a signal-bearing beam passes through a turbulent atmospheric channel, the atmospheric turbulence causes the optical power within the beam cross section to fluctuate randomly, as well as producing light-intensity scintillation, beam drift, and phase undulation. This reduces the communication performance of the optical communication system.

Figure 15.3 shows the impact of the atmospheric channel on the optical communication. Light-intensity flickering, beam drift, beam expansion, and phase fluctuation cause the optical signal to fluctuate randomly, which attenuates the optical power, decreases the signal-to-noise ratio, and reduces the communication quality. Considering the series of effects of atmospheric turbulence on optical wireless communication systems, it is important to study methods for suppressing atmospheric turbulence.

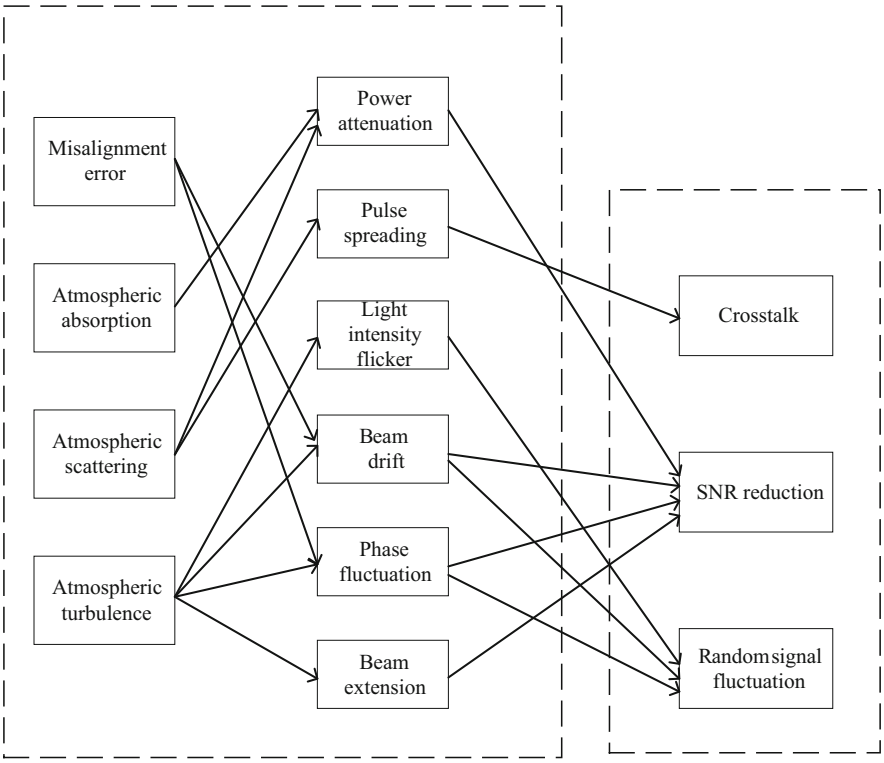


Fig. 15.3 Impact of the atmospheric channel on optical communication [6]

When a light beam passes through an atmospheric-turbulence channel, the light signal will be disturbed by turbulence, changing or even reducing the light power and resulting in the degradation of communication quality. Different atmospheric-turbulence scales will cause the light signal and beam quality to change differently; hence, the impact on the optical wireless communication system also differs. It has been shown that changes in the atmospheric-turbulence scale can have significant effects on optical wireless communication systems, including changes in the optical signal and beam quality, which, in turn, depend on the diameter of the laser beam and the geometry of the turbulence.

Light-Intensity Scintillation

When the cross-sectional diameter of the emitted beam is much larger than the geometric turbulence scale, the beam cross section includes many small-scale turbulence pairs, which causes the beam to have different effects, such as beam scattering and diffraction. At this time, the light intensity and phase will produce uncertain fluctuations, which affect the optical communication system performance and communication quality. This phenomenon is also known as light-intensity scintillation.

Beam Drift

As the beam passes through turbulence, it refracts. Multiple refractions will cause the beam to drift, and eventually lead to the beam not being irradiated vertically on the receiving surface. Beam drift causes random phase fluctuations in the transmitted optical signal, which greatly reduces the communication quality of the optical wireless communication system.

Beam Expansion

Atmospheric turbulence makes the atmospheric refractive index uncertain. This uncertain change and randomness expands the size of the beam in the transmission, resulting in the transmission of a beam whose cross-sectional energy is greatly reduced. When the energy of the beam is reduced to a certain degree, it cannot pass smoothly through the receiving lens to reach the photodetector. This reduces the communication quality of optical wireless communication.

Phase Fluctuation

In an atmospheric channel, the inhomogeneity of the transmission medium distorts the wavefront. This distortion makes it so the beam arrival angle does not match original arrival angle. It is also known as the arrival-angle fluctuation phenomenon. The wavefront of an optical signal in a uniform-medium transmission will not be distorted. Atmospheric turbulence can cause the distortion of the beam wavefront, it is called phase fluctuation.

Key Technologies for Suppressing Atmospheric Turbulence

Several technologies have been proposed to reduce the impact of atmospheric turbulence on optical wireless communication, and further improve the system performance and communication quality of optical wireless communication. Solutions and related technologies for suppressing atmospheric turbulence proposed by domestic and foreign scholars include large-aperture reception, adaptive optics, diversity, and partially coherent optical technologies. This section introduces several key technological principles for suppressing atmospheric turbulence and their research status.

Large-Aperture Receiving Technology

Large-aperture reception technology uses a large-diameter antenna or receiver to receive wireless signals. In the communication field, the main function of large-aperture receiving technology is to receive and process signals from different directions. Using a large-aperture antenna or receiver, the strength and accuracy of the received signals can be increased, and the quality and reliability of communications can be improved. Large-aperture receiving technology can be used in a variety

of applications, such as television and radio transmission, radar systems, satellite communications, and wireless laser communications.

Current Research Status

In 1973, Kerr et al. [7] showed that, under the same turbulence conditions, the fluctuation of the received optical power decreased as the receiving aperture increased, whereas when the emitted aperture was kept constant, the fluctuation of the received optical power decreased as the receiving aperture increased.

In 2006, Yuksel et al. [8] proposed an aperture-optimization method for optical antenna receivers that provided an optimal aperture size. They proved that the aperture-smoothing factor was independent of the aperture shape and was only related to the aperture area.

In 2008, Chen et al. [9] conducted a 16 km laser-communication link experiment to analyze the effect of atmospheric turbulence on the receiving system. They found that reception through a large-aperture antenna reduced turbulence-induced light-intensity fluctuations.

In 2009, Prokes [10] found that optical signals in optical wireless communication systems were affected by atmospheric turbulence as the communication distance increased. The scintillation effect of atmospheric turbulence could be suppressed by appropriately increasing the receiving-aperture size by using large-aperture receiving technology.

In 2009, Chen et al. [11] analyzed the relationship between the aperture-smoothing factor and the variance of light-intensity scintillation in a large-aperture receiving system. They determined the effect of different receiving-aperture sizes on the signal-to-noise ratio of the system with a given bit-error rate (BER), which laid the foundation for selecting the appropriate receiving-aperture size.

In 2011, Lou et al. [12] conducted a near-earth laser-transmission experiment and analyzed the effect of different receiving-aperture sizes on the optical-intensity fluctuation. The results showed that large-aperture receiving could receive a larger spot area, which effectively suppressed the optical-intensity scintillation effect of turbulence.

In 2012, Lu et al. [13] analyzed the effect of the antenna aperture on the quality of free-space optical communication and found that the signal scintillation index decreased with an increase in the receiving aperture, and the fading probability of the communication link similarly decreased.

In 2014, Wang [14] conducted experiments on atmospheric laser communication using different receiving-aperture sizes. They found that as the receiving-aperture size increased, the BER of the communication system decreased, suppressing the light-intensity fluctuation effect and improving the performance of the entire communication system.

In 2017, Li et al. [15] found that the scintillation index increased with an increase in turbulence intensity and conducted experiments on communication links at different turbulence intensities. They found that the scintillation index could be reduced with an increase in the receiving-aperture size, the scintillation index

could reach the saturation state, and an appropriate increase in the receiving-aperture size could effectively suppress the turbulence-scintillation effect.

In 2019, Soni et al. [16] studied the performance of a free-space optical communication system under different rainfall intensities using large-aperture reception and analyzed the suppression of atmospheric turbulence by the aperture-smoothing effect. This led to an average signal-to-noise ratio improvement of 1.58 dB using large-aperture reception.

In 2021, Dai [17] analyzed the effect of different aperture sizes on the light-intensity scintillation index in a wireless laser communication system using a large-aperture Cassegrain antenna as the receiving system. It was found that the light-intensity scintillation index of the large-aperture reception was small and stable, effectively suppressing the scintillation effect of atmospheric turbulence.

Suppression Principle

In optical wireless communication systems, because atmospheric turbulence caused by light-intensity flicker will lead to received optical-power fluctuation at the receiving end, the antenna size at the receiving end can be appropriately increased to reduce the optical power of the fluctuation variance. Large-aperture receiving technology increases the size of the receiving end of the aperture, which allows the receiving end to receive more light energy to inhibit the effect of atmospheric turbulence.

Large-aperture reception technology is widely used in optical wireless communication systems, and the suppression of the scintillation effect is particularly obvious. Large-aperture receiving technology reduces the flickering effect caused by atmospheric turbulence by increasing the receiving aperture at the receiver end; this phenomenon is known as the aperture-smoothing effect. Under different atmospheric-turbulence intensities, different receiving-aperture sizes have different effects on the scintillation effect. Therefore, the aperture-smoothing factor characterizes the aperture-smoothing effect. Considering the aperture-smoothing factor under different turbulence conditions, an optimal aperture size of the receiving antenna can be found to suppress the turbulence scintillation effect [18].

The Cassegrain-telescope receiving system is a typical and commonly used large-aperture receiving system. The Cassegrain telescope system is comprised of two mirrors. As shown in Fig. 15.4, the incident laser beam passes through the first primary reflector with a parabolic surface pattern and is then output by the secondary reflector with a hyperbolic surface pattern. The power of the received optical signal can be increased by appropriately increasing the aperture of the Cassegrain telescope. This also has an aperture-averaging effect, weakens the fluctuation, and effectively suppresses the effect of atmospheric turbulence on the beam.

To a certain extent, large-aperture receiving technology suppresses the impact of light-intensity scintillation on the optical communication link; however, the aperture size cannot increase infinitely. An increased aperture means an increase in the size and weight of the receiver; thus, applications with high portability requirements will be limited. Moreover, it increases the design difficulty of other parts of the system to a certain extent.

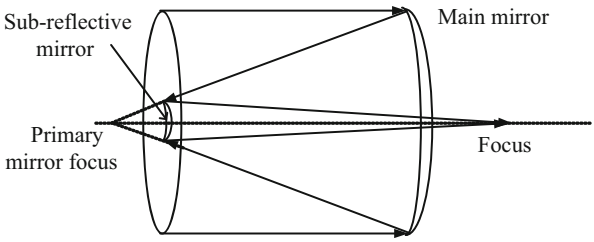


Fig. 15.4 Schematic diagram of a large-aperture Cassegrain telescope [19]

Partially Coherent Optical Technology

Partially coherent optical-transmission techniques have been applied in optical communication. The phase and amplitude of the signal may change, owing to atmospheric turbulence, and decrease the signal coherence. In a partially coherent optical-transmission technique, the phase and amplitude of the optical signal can change to some extent while still maintaining a certain degree of coherence. Partially coherent optical-transmission technology can process and repair signals using a series of modulation and demodulation techniques, thereby improving the transmission reliability and performance.

Current Research Status

In 1978, Wu et al. [20] discovered that partially coherent beams had a strong antiturbulence effect by analyzing the transmission characteristics of partially coherent Gaussian–Schell-model (GSM) beams under atmospheric turbulence.

In 2003, Ricklin et al. [21] investigated whether the BER of free-space optical communication systems was affected by the turbulence strength and degree of beam coherence. The final results demonstrated that reducing the beam coherence led to a reduction in the degree of optical power jitter at the receiver, which led to a reduction in the BER of the communication system.

In 2003, Aristide et al. [22] used partially coherent light for transmission in free-space optical communication systems. They showed that partially coherent light was less affected by atmospheric turbulence than fully coherent light, and the effect of turbulence could be better suppressed by appropriately reducing the beam coherence.

In 2010, Borah et al. [23] investigated optimizing the coherence length of a spatially partially coherent beam for free-space optical communication under weak turbulence conditions. The effects of different parameters on the coherence length were discussed. Optimizing the coherence length reduced the turbulence effect and improved the system performance.

In 2011, Drexler et al. [24] improved the performance of free-space optical communication systems by controlling the coherence of the transmitted beams. It was demonstrated that the use of partially coherent beams reduced the scintillation values by 50%, compared to fully coherent beams under different turbulence-intensity conditions.

In 2014, Zhang et al. [25] investigated the transmission characteristics of partially coherent Gaussian beams with different coherences and experimentally derived the scintillation index for beam transmission in atmospheric turbulence. The results showed that reducing the spatial coherence of the beam could suppress the light-intensity scintillation in turbulence.

In 2016, Zhang et al. [26] derived the Rayleigh-interval and turbulence-distance analytic equations for a radial partially coherent GSM array beam during atmospheric-turbulence transmissions. They showed that the atmospheric-turbulence effect could be effectively suppressed by appropriately reducing the beam coherence.

In 2019, Zhang [27] investigated whether anisotropic atmospheric turbulence could lead to poor optical-signal quality. They proposed a novel twisted rectangular Laguerre–Gaussian–Schell model (TRLGCSM) beam generated by partially coherent light superposition and verified that the beam could effectively suppress the effects of atmospheric turbulence.

In 2019, Yu et al. [28] found that part of a coherent beam was affected by atmospheric turbulence. The beam-drift phenomenon caused by atmospheric turbulence could be reduced by changing the coherent structure of the optical field and reducing the beam coherence.

In 2019, Zhang [29] conducted indoor simulation experiments using a liquid-crystal spatial light modulator (LC-SLM) to generate a partially coherent beam. Using the LC-SLM for atmospheric-turbulence simulation, they verified that the partially coherent beam generated by the LC-SLM had better turbulence resistance than a fully coherent beam.

In 2022, Yang et al. [30] analyzed the transmission characteristics of a partially coherent radially polarized twisted beam under atmospheric turbulence and derived its twist factor. They showed that the larger the twist factor of the beam, the better the resistance to turbulence.

In recent years, Xi'an University of Technology has conducted extensive research on partially coherent light transmissions in the atmosphere [31]. In 2014, the transport properties of partially coherent GSM beams in atmospheric turbulence were studied by Han et al. [32, 33]. They verified that partially coherent light had better turbulence resistance than fully coherent light. The experimental results showed that the beam was less affected by beam drift when its spatial coherence length was greater than 0.005 m within a transmission distance of 10 km.

In 2015, Zhang et al. [34] studied the scintillation effect of partially coherent light subjected to atmospheric turbulence and analyzed the variance in the light-intensity fluctuation of partially coherent light under different atmospheric-turbulence intensities. The experiments showed that partially coherent light had a better anti-turbulence effect under the same turbulent conditions, and that the turbulence effect could be effectively suppressed by reducing the coherence of the beam.

In the same year, Wang et al. [35, 36] studied beam-expansion expressions for partially coherent GSM beams and discussed the transmission characteristics when the beams were transmitted obliquely and horizontally. The light-intensity distribution when the beam was normalized for a horizontal transmission was the same as

that for an oblique-range transmission. The beam was less affected by turbulence during the oblique-range transmission.

Wang et al. [37] studied the transmission characteristics of a partially coherent electromagnetic Schell-model beam in upstream and downstream links in atmospheric turbulence. They demonstrated that the beam could transmit over longer distances in the downlink transmission.

In 2016, Wang et al. [38] studied the transmission characteristics of partially coherent vortex beams under atmospheric turbulence. It was demonstrated that the partial vortex beam had a better transmission performance than the full vortex beam. Zhang et al. [39] studied the transmission characteristics of a partially coherent GSM array beam under atmospheric turbulence. It was demonstrated that the array beam had a better turbulence-suppression ability than a single beam. Zhang et al. [40] studied and analyzed a partially coherent Gaussian–Schell-model pulsed beam (GSMP) and verified that its expansion was less affected by turbulence than that of a fully coherent beam.

In 2017, Wang et al. [41] derived an expression for the average scattering radius of a GSM beam at the receiving end. It was verified that the larger the initial radius of the GSM beam or the smaller the coherence, the less its average scattering radius was affected by atmospheric turbulence. In the same year, Xue et al. [42, 43] investigated the transmission characteristics of GSM beams at various turbulence scales and intensities. This study demonstrated the ability of an inclined-range transport to resist turbulence better than a horizontal transport under the same turbulence conditions.

Wang et al. [44] studied the transmission characteristics of partially coherent Airy beams. It was shown that the smaller the beam truncation factor, the less it was affected by turbulence and the more complete the light-intensity distribution. Zhang et al. [45] studied the transmission characteristics of partially coherent Airy beams and showed that they had less beam drift than Gaussian beams under the same turbulent conditions. Moreover, they had strong turbulence-suppression and recovery properties. Wang et al. [46] studied the transmission characteristics of partially coherent off-axis vortex beams and derived an expression for the light-intensity distribution.

In 2019, Wang et al. [47, 48] investigated a partially coherent electromagnetic Gaussian–Schell-model vortex (EGSMV) beam and verified that the larger the topological charge of this beam in an atmospheric-turbulence transmission, the more information the receiver side received, and the larger the midline dark spot of the beam.

Suppression Principle

It is commonly believed that an ideal point light source emits perfectly coherent light; however, as the laser propagates, its phase and amplitude fluctuate to varying degrees, and the coherence degree of the perfectly coherent light is reduced. Therefore, a fully coherent light source does not exist and partially coherent light, which lies between fully coherent and incoherent light, is prevalent.

Highly coherent laser light is prone to scattering in inhomogeneous media, especially in the atmosphere, where the transmission is affected by turbulence, resulting in an uneven light-intensity distribution. Therefore, the atmospheric-turbulence effect can be suppressed by appropriately reducing the optical coherence. Studies have shown that partially coherent light is less affected by turbulence effects in atmospheric transmission, its scintillation index is less than that of highly coherent light, and it has better turbulence resistance [49].

Partially coherent light can suppress turbulence because it has multicolor properties. Even if it is affected by atmospheric turbulence, its transmission effect is better than that of single-frequency coherent light, so that transmissions using partially coherent light have less scintillation, and more light energy can be received at the receiving end. However, coherent light transmission in turbulence causes the beam to refract in all directions; therefore, only part of the light energy is received at the receiving end. A schematic diagram of partially coherent–light turbulence suppression is shown in Fig. 15.5.

The GSM beam [50] is a typical partially coherent light beam with a spatial correlation structure. Regarding the evolutionary properties of this beam in atmospheric turbulence, it was found that the GSM beam had better turbulence resistance compared to a fully coherent beam. In addition, some partially coherent beams have other novel correlation structures, such as Hermite–Gaussian correlated Schell-Model (HGSM) beams [51], Cosine-Gaussian–Schell-Model (CGSM) beam arrays [52], and Multi-Gaussian–Schell-Model (MGSM) beams [53]. These beams can effectively suppress the influence of negative atmospheric-turbulence effects when transmitted through atmospheric turbulence. However, for some coherent light beams in the transmission process, the beam-divergence angle is larger, resulting in a correspondingly larger signal-energy loss.

Diversity Technology

Diversity is a technique for receiving signals in parallel through multiple independent receiving channels, and then combining these signals for processing. Its main

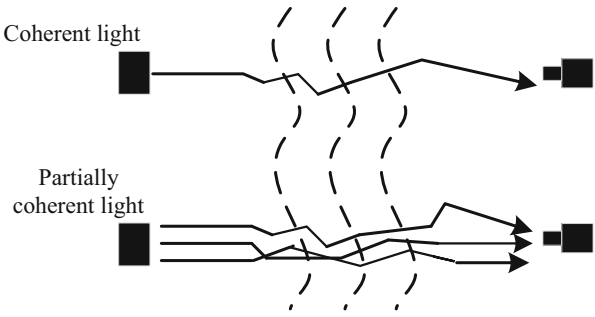


Fig. 15.5 Schematic diagram of partially coherent optical turbulence suppression [17]

purpose is to improve the reliability and performance of the signal reception. In communication systems, the diversity technique is commonly used to reduce the influence of multipath fading, interference, and noise on signal transmissions, thereby improving the signal quality and fault tolerance of the system. It reduces the signal distortion and degradation of the signal-to-noise ratio by simultaneously receiving multiple independent signal paths, and then processing and combining them.

Current Research Status

In 1996, Ibrahim [54] hypothesized that the output signals of small-aperture receivers faded independently of each other and investigated the spatial-reception diversity technique using multiple reception apertures.

In 2004, Pan et al. [55] analyzed the optical-intensity flicker characteristics of a transmission link using a multibeam transmission technique for a ground-to-ground optical communication system. They found that the multibeam transmission technique could effectively suppress the turbulence-induced optical-intensity flicker in the entire turbulence region from weak to strong.

In 2005, Trisno et al. [56] used a time-delay diversity technique to suppress the fading of optical signals due to atmospheric turbulence and obtained signal-to-noise gains of 3.8 dB, 4.3 dB, and 4.4 dB, by increasing the time delay.

In 2006, Anguita et al. [57] used a multibeam transmission technique, analyzed two space-time coding modes, and demonstrated that both modes had higher signal-to-noise gains compared to single-beam transmissions under the influence of atmospheric turbulence.

In 2008, Neda et al. [58] analyzed a ground-based optical-flux system using multibeam emissions and multiple avalanche-photodiode detection reception. They discussed the gain-size selection for merging multiple avalanche-photodiode output signals.

In 2011, Ke et al. [59] used a diversity technique at the receiver side, and the light-intensity flicker effect caused by atmospheric turbulence was effectively suppressed by multiaperture reception. They verified that the BER and aperture-smoothing factor of the system were further reduced with an increase in the number of reception apertures.

In 2012, Ke et al. [60] analyzed the light-intensity probability distribution of signals transmitted through multiple beams under different turbulent environments. The results showed that multiple-beam transmissions could reduce the BER of the system, decrease the light-intensity fluctuation effect, and improve the performance of the communication system under the same turbulent environment.

In 2014, Long [61] demonstrated that aperture diversity could reduce the optical-intensity flicker index and that the system performance was better than that of a large-aperture receiving system. They used a diversity technique and numerical simulations with a phase screen and compared it with a large-aperture receiving system.

In 2015, Kaur et al. [62] investigated the combined effect of atmospheric turbulence on optical wireless multiple-input multiple-output (MIMO) communication

systems under the influence of different weather conditions and compared the BER and outage capacity using aperture-averaging techniques under the same conditions. The results showed that MIMO techniques under different weather conditions effectively suppressed atmospheric turbulence.

In 2016, Bhatnagar et al. [63] studied the effect of alignment errors on an optical MIMO communication system under a gamma–gamma turbulence channel model. They analyzed the performance of a communication system under equal-gain merging and maximum-ratio merging and concluded that the maximum-ratio merging criterion could suppress alignment errors.

In 2017, Meng [64] investigated optical wireless communication systems using the split-set merging technique and showed through experimental simulations that the light-intensity flicker effect caused by atmospheric turbulence could be most effectively suppressed using the equal-gain merging technique.

In 2019, Kumar et al. [65] analyzed the performance of optical wireless MIMO systems under gamma–gamma turbulent channels using different diversity techniques. The results showed that the optical wireless MIMO system could reduce the system BER and improve its reliability under a gamma–gamma turbulent channel.

In 2019, Xu [66] used a diversity technique to derive the probability distribution of light intensity under multibeam emissions. Improved simulation experiments showed that under the same turbulent conditions, it had the right antiturbulence ability, compared to single-aperture emission, which effectively mitigated the light-intensity scintillation effect.

In 2021, Ding [67] analyzed the variation in the receiving-system performance by using a spatial-diversity technique for a partially coherent optical transmission system. The experimental results showed that the signal-to-noise ratio of the communication system could be improved by using a diversity technique, and the performance of the communication system was improved under the influence of turbulence.

In 2021, Dai [17] demonstrated that the scintillation effect could be suppressed by increasing the number of receiving apertures in a large-aperture Cassegrain receiving system. They reduced the light-intensity scintillation index through a diversity technique.

In 2022, Liu [68] investigated optical wireless MIMO communication technology, and the atmospheric-transmission characteristics of a single-input single-output (SISO) system and a MIMO system were comparatively studied using liquid-crystal spatial light modulators to simulate atmospheric turbulence. It was demonstrated that the MIMO communication system had a lower link cost and improved communication performance compared to the SISO communication system.

Suppression Principle

The application of diversity technology in optical wireless communication systems involves increasing the number of antennas at the transmitter and receiver ends. A block diagram of the system is shown in Fig. 15.6. The transmitter end of the system outputs the transmitted laser beam using multiple transmitting antennas. When it

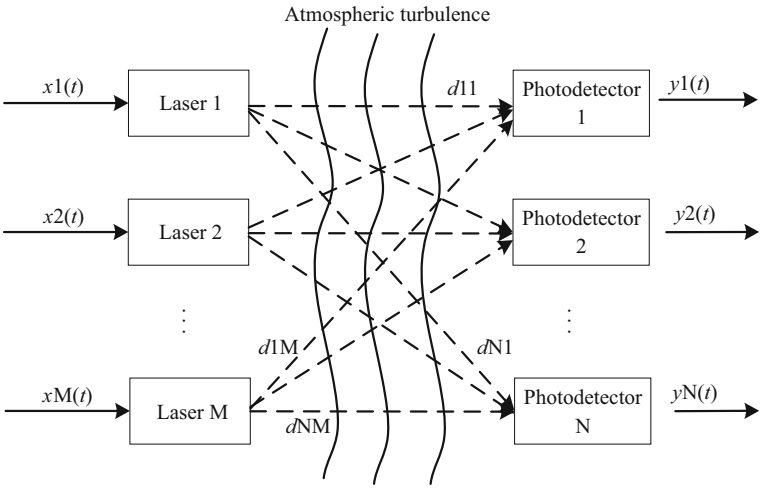


Fig. 15.6 Block diagram of the diversity system [68]

reaches the receiver end, it is received by multiple receiving antennas. The atmospheric-channel attenuation experienced by each signal is unrelated, which can effectively suppress the atmospheric-turbulence effect and improve the communication quality of the entire communication system.

Diversity technology includes transmit and receive diversity. Transmit diversity often uses multibeam transmission technology, which is a free-space laser transmission technology that can effectively suppress the scintillation effect by transmitting two or more superpositioned independent and incoherent lasers through different paths to the receiving end. This smooths the light-intensity scintillation on the far-field receiving plane.

For receiving diversity, antenna-array receiving technology is more commonly used, by setting multiple independent small-aperture receiver antenna arrays instead of large aperture receivers. Under the same receiving-area conditions, the method can receive multipath signals to a greater extent, reduce the impact of turbulent channels, and improve the reliability of the communication system. The structure is simple, flexible, and has a lower cost.

The receiver side combines the multiple received fading signals to obtain a gain signal, which improves the quality of the received signal. Therefore, the merging technique is the most important aspect of the diversity-reception technique. The most commonly used diversity-set merging techniques are maximum ratio, selection, and equal-gain combining [69].

Selection combining is a relatively simple merging technique, in which each branch is demodulated so that their signal-to-noise ratios are the same; then, a branch is selected at the receiving end for demodulation. This approach is difficult to implement in practice because the signal-to-noise ratio of the communication system is not easily measured.

Maximum-ratio combining uses the ratio of the voltage value of each branch signal and the noise power as the weight; then, at the receiving end of each signal, the weight is used for phase correction and according to the gain coefficient to make each signal sum. Finally, it is received in the detector for coherent detection. This reception method is not ideal. The signal can also be demodulated through the maximum-ratio merging method.

Equal-gain combining does not require the weighting of each signal. It directly assumes that each path has the same weight and the same signal is directly superimposed. The results obtained using equal-gain combining and maximum-ratio combining are similar.

The improvements in the gain and signal-to-noise ratio of the optical communication system differ among the three merging methods. Under the same experimental conditions, the improvement in system performance was most evident for maximum-ratio merging, followed by equal-gain merging, and selective merging had the least improvement in system performance.

Adaptive-Optics Technology

Adaptive optics adjusts the shape of optical elements and surface topography in real time, according to changes in the optical system. It is primarily used in astronomy, optical communication, laser processing, and other fields to improve the performance and effects of optical systems. Adaptive-optics technology can effectively improve the performance of optical systems, especially in environments with atmospheric turbulence, significantly reduce the aberration and blurring phenomena caused by atmospheric turbulence, and improve the resolution and transmission distance of astronomical observations and laser communication.

Current Research Status

In 1996, Tyson et al. [70] concluded from a turbulence-theory analysis that conventional adaptive-optics (AO) techniques were effective in reducing the optical-signal attenuation and fluctuations in the uplink and downlink of a ground-space free-space optical communication system under weak-turbulence conditions.

In 1998, Roggemann et al. [71] used adaptive optics and proposed the joint use of two deformation mirrors to correct the light-intensity scintillation effect caused by atmospheric turbulence in laser atmospheric transmissions.

In 1998, Barbier et al. [72] used adaptive-optics techniques to simulate the performance of communication systems using the phase-screen method. The effects of the dynamic range and adaptive-optical aperture size of the wavefront sensor on the mean and variance of the optical intensity were analyzed.

In 2001, Li et al. [73] proposed a correction method using two adaptive-optics systems in series. The control performance and noise-transfer characteristics of this approach were analyzed. The simulation results showed that the tandem structure could substantially improve the utility of the adaptive-optics system in controlling the wavefront aberrations of atmospheric turbulence.

In 2004, Thomas et al. [74] used adaptive optics to compensate for the tilt aberration of the wavefront using a tilt mirror and the higher-order aberration using a micro-electromechanical system (MEMS) deformation mirror. Using this dual-deformation-mirror correction method resulted in a twofold increase in the received signal intensity compared to when only the tilt aberration was corrected.

In 2010, Xia et al. [75] used an adaptive-optics technique to effectively correct the aberrated wavefront and experimentally demonstrated that this technique could improve the communication system gain, reduce the system transmit power, and lower the BER under the same BER conditions.

In 2011, Han [76] used adaptive optics to compensate for free-space optical communication systems and provided the corresponding expressions for the BER and coupling efficiency, as well as the relationship between the coupling efficiency and Strehl ratio.

In 2013, Wu [77] verified through simulation experiments that adaptive optics could effectively correct an aberrated wavefront and improve the phase matching of the local-oscillation light and signal light to reduce the system BER.

In 2014, Hashmi et al. [78] used adaptive-optical techniques to simulate an interstellar optical communication system under moderate background optical noise and strong turbulence. The results showed that the Strehl ratio of the system improved from 0.30 to 0.75 after reaching a steady state with adaptive-optical closed-loop control.

In 2015, Cvijetic et al. [79] conducted experiments using adaptive optics on coherent optical communication systems under the influence of atmospheric turbulence in the ocean. The results showed that the adaptive-optics technique could compensate for the phase distortion of the optical signal and reduce the BER of the communication system.

In 2015, Wright et al. [80] used adaptive optics to build an optical-communication telescope ground-observation communication system and corrected the aberrated wavefront, resulting in an improved Strehl ratio of the received beam from 0.03 to 0.66 after correction.

In 2017, Escarate et al. [81] designed a model-predictive controller to calibrate an adaptive-optics system with 140 units and two deformation mirrors spaced 3 m apart. The results showed that this method could effectively correct the amplitude and phase of the laser beam.

In 2018, Li [82] studied several methods to suppress atmospheric turbulence and completed a wavefront detection-free adaptive-optical atmospheric-turbulence suppression experiment using the stochastic parallel gradient-descent (SPGD) algorithm. The results showed that the corrected fiber-coupling efficiency could be improved from 2.7 dB to 9 dB at different turbulence intensities.

In 2019, Lin et al. [83] predictively estimated state variables for adaptive-optical systems and implemented closed-loop correction using a linear quadratic Gaussian control technique based on state regulation. The results showed that the method could better improve the system response speed and stability compared to a proportional–integral–derivative (PID) control algorithm.

In 2019, Chen et al. [84] used an adaptive-optics technique to apply a continuous-deformation mirror with 137 cells to an adaptive-optics system, which gave a BER of less than 10^{-6} for the communication system by correcting the aberrated wavefront. It was significantly reduced compared to the uncorrected BER.

In 2020, Banet et al. [85] proposed a method to compensate for the beacon using adaptive optics. Under the same turbulent conditions, compared to the uncompensated fiducial adaptive-optics system, the Shack–Hartmann wavefront sensor of the adaptive-optics system with compensated beacons had a 17% performance improvement.

In 2022, Yang et al. [86] analyzed a K-L-mode SPGD algorithm using adaptive optics to correct an optical-wavefront aberration under the influence of atmospheric turbulence. They verified that the convergence speed of the SPGD algorithm in this mode had improved by 47.5%, compared to the conventional SPGD algorithm.

In 2023, Li et al. [87] investigated free-space optical communication under the influence of atmospheric turbulence and proposed an advanced multifeedback SPGD (AMF-SPGD) algorithm to enhance the wavefront-aberration correction performance. They verified that the algorithm could improve the Strehl ratio and reduce the BER of the communication system.

Recently, Xi'an University of Technology has conducted research on many aspects of adaptive optics [88]. In 2018, Wu et al. [89] corrected a distorted wavefront using adaptive optics without wavefront detection, based on the SPGD algorithm. The corrected Strehl ratio was improved from 0.15 to 0.81, which resulted in a larger received spot and mitigated the effect of atmospheric turbulence on the wavefront distortion.

In the same year, Wang et al. [90, 91] used adaptive optics without wavefront detection based on the Gerchberg–Saxton (GS) and SPGD algorithms for wavefront correction of vortex beams. The simulation results showed that the GS algorithm outperformed the SPGD algorithm in terms of correction results.

In 2019, Ke et al. [92] used an adaptive-optics technique with wavefront detection based on double fuzzy control to correct an aberrated wavefront. The simulation results showed that this method could effectively correct the wavefront. In the same year, Yang et al. [93, 94] used adaptive optics with wavefront detection and effectively corrected aberrated wavefronts using computational optical-field imaging.

Zhang et al. [95] proposed an improved SPGD algorithm that used wavefront-sensor-free detection adaptive optics. By parallelizing the SPGD algorithm, the experimental results showed that this method could improve the algorithm-convergence and wavefront-correction speeds. Han et al. [96] used two liquid-crystal spatial light modulators for wavefront-aberration simulation and wavefront correction according to the principle of adaptive optics. The corrected Strehl ratio was improved from 0.38 to 0.82, and the method reduced the experimental cost.

In 2020, Li [97] used adaptive optics without wavefront detection and investigated a wavefront-correction method using the deformation-mirror eigenmode. The Strehl ratio of the correction result improved from 0.56 to 0.88, verifying that the

method effectively corrected wavefront aberrations caused by atmospheric turbulence and that the algorithm converged quickly.

In the same year, Cui [98] used a phase-diversity (PD)-based wavefront detection-free adaptive-optical technique for correcting vortex beams. The experimental results showed that this method could improve the signal power, reduce inter-code crosstalk, and effectively mitigate the turbulence effect.

In 2021, Cheng [99] used a PID control algorithm and a large-amplitude aberration wavefront residual-correction algorithm to correct the problem of insufficient range of the deformation mirror in an adaptive-optics system. The experimental results showed that the large-amplitude aberration wavefront could be effectively corrected using the residual correction algorithm.

Suppression Principle

Adaptive-optics technology can automatically improve the beam-wavefront quality. Therefore, the application of adaptive optics to optical wireless communication systems has great potential for suppressing the effects of atmospheric turbulence. Adaptive-optics systems are divided into two categories: those with wavefront sensors [100] and those without wavefront sensors [101].

Adaptive-Optics Systems with Wavefront Sensors

The adaptive-optics system with a wavefront sensor is shown in Fig. 15.7 and consists of three main components: a wavefront sensor, a wavefront controller, and a wavefront corrector. The general correction process is as follows: The emitted beam is affected by atmospheric turbulence, and the beam wavefront is distorted. The distorted beam passes through the corrector for the first time to collect information and then arrives at the beam splitter. Part of the beam is received, and part of the beam is transmitted to the wavefront sensor to measure the beam-wavefront information. This information is converted into electrical signals that are sent to the wavefront controller for wavefront reconstruction. A suitable control algorithm is selected to generate the control signal. Finally, the wavefront corrector receives the control signal, allowing the corrector to correct the distorted wavefront, forming a closed-loop system.

Wavefront sensors are one of the main components in adaptive-optics systems. Their role is to measure aberrated wavefront information in real time and send it to the wavefront controller. Commonly used wavefront sensors are transverse-shear interferometers [102], Shack–Hartmann sensors [103], curvature wavefront sensors [104], and point-diffraction interferometer sensors [105]. The wavefront controller processes the wavefront information from the wavefront sensor into a control signal and performs two functions: wavefront reconstruction and control-signal generation. Wavefront correctors are essential components of adaptive-optics systems to correct wavefront errors in a timely manner. Common wavefront correctors include liquid-crystal light modulators (LC-SLM) [106] and deformable mirrors (DM) [107].

In adaptive-optics systems with wavefront sensors, the correction accuracy of the distorted wavefronts varies between different control algorithms; commonly used ones are classic and iterative control algorithms. Among the classic control

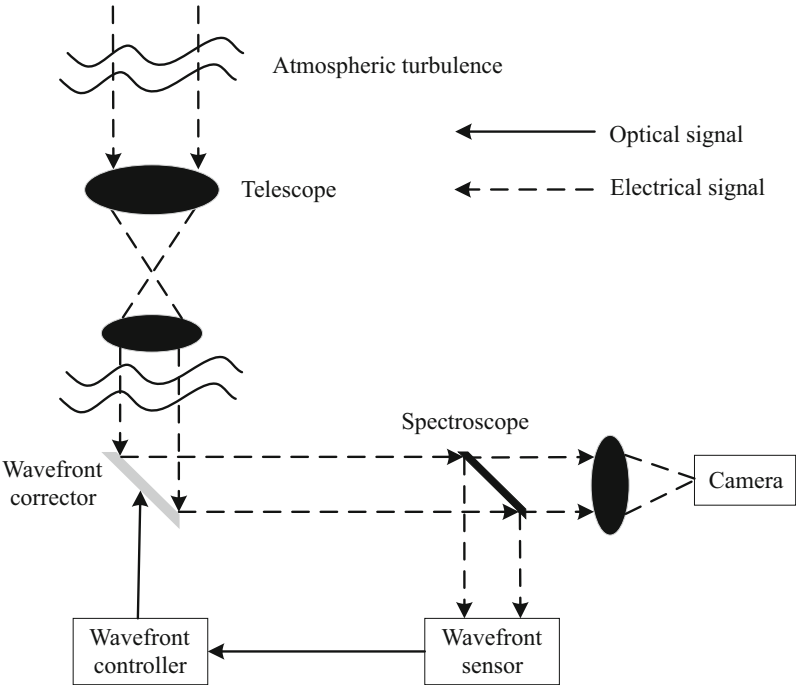


Fig. 15.7 Adaptive-optical system with wavefront sensor [88]

algorithms, the PID control algorithm [108] is widely used in adaptive optics as the most common algorithm.

The main principle of this control algorithm is as follows: First, the command matrix of the system is determined using the push–pull method. Subsequently, the command matrix receives the error slope and converts it into an error-voltage signal. The obtained voltage is then sent to the deformation mirror for wavefront correction by the PID control system. Finally, the wavefront sensor collects the wavefront slope again to achieve closed-loop control.

The Gauss–Seidel iterative algorithm is widely used in adaptive optics to correct aberrated wavefronts, and its effectiveness in adaptive optics is related to the accuracy of the iterative error. The Gauss–Seidel iterative algorithm [108] is based on calculating the information collected by the wavefront sensor, which is then repeated in the iterative formula until the value of the iterative judgment condition is less than the iterative accuracy error; otherwise, the feedback is repeated. Subsequently, the system can proceed to the next slope-acquisition process and repeat the above steps to complete the closed loop. The iterative learning-control (ILC) algorithm [108] is a new iterative algorithm that establishes a relationship between the Zernike polynomial [109] and the control voltage signal of the calibrator to correct the wavefront.

However, the iterative algorithm suffers from problems, such as the range of values used for the parameters. Therefore, its practical applications are limited. In contrast, the PID control algorithm is simple in principle and has better overall performance.

Adaptive-Optics Systems Without Wavefront Sensors

Figure 15.8 shows an adaptive-optics system that does not require a wavefront sensor for wavefront sensing and reconstruction. The wavefront sensor-free system consists of an imaging detector, a wavefront corrector, and a wavefront controller.

The calibration process is as follows: The wavefront calibrator first calibrates the received beam. Then, the beam is projected incident on the imaging detector to collect the information and transmit it to the controller. The controller calculates the voltage signal required by the calibrator using a suitable control algorithm and iterates until calibration is achieved. Adaptive-optical intelligent optimization algorithms for wavefront sensor-free detection fall into two main categories: mode-free and mode optimization methods.

The mode-optimization method involves selecting a suitable base mode to describe the wavefront error and correcting the wavefront distortion by optimizing the base = mode coefficients. Mode-free optimization directly controls a deformable mirror to correct wavefront distortions and optimize system performance-evaluation metrics. Currently, the SPGD algorithm [110], simulated-annealing (SA) algorithm [111], and genetic algorithm (GA) [112] are commonly used among pattern-free optimization algorithms. The SPGD algorithm can estimate the gradient, has fast convergence, is more stable, has simple parameters, and is easy to implement.

The basic working principle of the SPGD algorithm [89] is to estimate the control parameters based on the change values of the performance-evaluation index, and to search for the control parameters in the gradient-descent direction using an iterative approach. It first generates the initial voltage signal of the deformed mirror and the

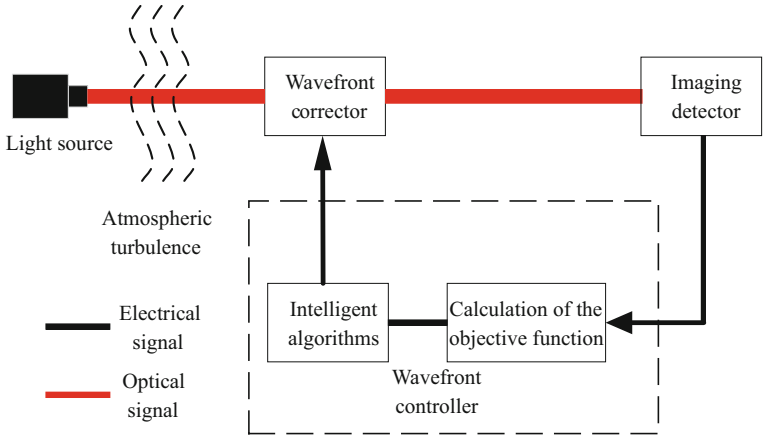


Fig. 15.8 Adaptive-optics system without wavefront detection [89]

random-disturbance voltage signal conforming to the Bernoulli distribution. Next, it adds a positive disturbance voltage and a negative disturbance voltage, respectively, and based on the initial voltage signal calculates the change in the system objective function under the voltage. Finally, it calculates the voltage signal to be loaded in this iteration according to the iterative formula. This is done until the target function satisfies the initial conditions set by the system, so that the system-evaluation index reaches the optimal value, thus, correcting the distorted wavefront.

Many researchers have proposed improved SPGD algorithms [87, 95], all of which can ultimately be very effective in improving the performance of communication systems and enhancing the convergence speed and correction effect of algorithms.

Development Trends

Optical wireless communication systems are susceptible to atmospheric turbulence, which distorts the transmitted beam wavefront and reduces the optical power of the transmitted beam, thereby increasing the BER of the communication system and reducing its quality. Researchers have adopted techniques, such as large-aperture reception, partially coherent light technology, diversity technology, and adaptive optics, to suppress the effects of atmospheric turbulence on optical wireless communication.

For large-aperture receiving technology, the size of the antenna at the receiving end can be appropriately increased to reduce the light-intensity flicker. Currently, many scholars use large-aperture receiving technology to suppress the flicker effect of turbulence, establish different smoothing-factor models according to the different turbulence intensities to design large-aperture receiving systems, and select the appropriate aperture to reduce the weight of the receiving end as much as possible.

Partially coherent light techniques are also highly effective in suppressing atmospheric turbulence, and it is important to investigate new types of partially coherent beams to suppress atmospheric turbulence.

Diversity is a fundamental technique for antifading in wireless communications [113] and can effectively suppress the fading caused by atmospheric turbulence. This technique is low-cost, technically simple, and easy to apply in engineering and has broad development prospects in the future.

Adaptive optics is one of the most widely used techniques for suppressing atmospheric turbulence. It can directly improve the quality of communication systems by correcting distorted wavefronts in real time. Various improvements have been proposed to improve the efficiency of single-mode fiber coupling through adaptive optics to improve communication quality. In recent years, multiple deformation mirrors have been used in adaptive-optics systems and suitable control algorithms have been optimized to improve calibration results. New techniques and methods have been used to improve the calibration of adaptive-optics systems, making them more valuable.

Summary and Outlook

This chapter analyzed domestic and foreign research on suppressing atmospheric-turbulence effects in optical wireless communication systems. The main technologies for suppressing atmospheric-turbulence effects were large-aperture receiving, partially coherent light, diversity, and adaptive-optics technology.

This chapter analyzed the principle of each technology for suppressing atmospheric turbulence. To a certain extent, these technologies can effectively suppress atmospheric turbulence, compensate for distorted wavefronts, and improve the performance of optical wireless communication systems. However, they are not perfect in engineering practice. Table 15.1 shows the advantages and disadvantages of each key technology.

Therefore, further research and discussion on methods to suppress atmospheric turbulence are of great importance.

Table 15.1 Advantages and disadvantages of key technologies for suppressing atmospheric turbulence

Related technology	Advantages	Disadvantages
Large-aperture receiving technology	An appropriate increase in the size of the receiving aperture reduces the variance of the optical-power fluctuations and effectively suppresses the light-intensity flicker effect of atmospheric turbulence	The increased aperture makes the receiver larger and heavier, making it less portable
Partially coherent light technology	Partially coherent light has good resistance to turbulence and effectively suppresses light-intensity flicker effects	It has a large beam-divergence angle and severe beam extension at the receiving end
Diversity technology	Multibeam emission with incoherent superposition at the receiving end smooths the optical power and effectively suppresses the light-intensity flicker caused by atmospheric turbulence	Multitransmitter and multireceiver diversity systems are complex and limited in their practicality
Adaptive optics with wavefront sensors	Automatically improves the optical-wavefront quality, enabling the effective correction of distorted wavefronts	When the communication distance is long, it is difficult for the wavefront sensor to measure the wavefront-aberration information, which makes it impossible to reconstruct the wavefront and compensate for the aberration in real time
Adaptive optics without wavefront sensors	Corrects distorted wavefronts without wavefront reconstruction	Higher cost, requiring additional equipment, and increased system complexity

1. Optical wireless MIMO technology can make full use of spatial spectrum resources and effectively suppress atmospheric-turbulence effects; however, this technology is not yet mature in the optical wireless field, and further research is necessary.
2. Partially coherent vortex beams have a better suppression effect on atmospheric turbulence, and further research on their generation is required.
3. For adaptive-optics technology, further study is needed on deformation mirrors and control algorithms, as well as using multiple deformation-mirror combinations for correction. Finally, the algorithm convergence should be studied to improve the accuracy of deformation-mirror correction.
4. In addition to the aforementioned techniques, new methods for suppressing atmospheric turbulence should be studied.

References

1. Ke Xizheng, Deng Lijun. optical wireless communication [M]. Beijing: Science Press, 2016.
2. Ke X Z, Ning C, Wang J. Crosstalk analysis of orbital angular momentum-multiplexed state under atmospheric turbulence [J]. Infrared and laser engineering, 2018, 47 (11): 402–408.
3. Zhang Ying, Ke Xizheng, Chen Mingsha. Simulation experiments on wavefront distortion correction for excited Brillouin scattering [J]. Infrared and Laser Engineering, 2018, 47(11): 395–401.
4. Hocking, W. K. Measurement of turbulent energy dissipation rates in the middle atmosphere by radar techniques [J]. Radio Science, 1985, 20(6), 1403–1422.
5. Meng Ce. Research on Signal Equalization Technology for optical wireless communication System at Receiver End [D]. Changchun: Changchun University of Technology, 2019.
6. Ke Xizheng, Wu Jiali, Yang Shangjun. Progress and prospects of atmospheric turbulence research for optical wireless communication [J]. Journal of Radio Wave Science, 2021, 36(03): 323–339.
7. Kerr R J, Dunphy R J. Experimental effects of finite transmitter apertures on scintillations* [J]. Journal of the Optical Society of America, 1973, 63(1): 1–8.
8. Heba Yuksel, Christopher C. Davi. Aperture averaging analysis and aperture shape invariance of received scintillation in free-space optical communication links [J]. Free-Space Laser Communications 2006, 63(4): 63041E.1–63041E.11.
9. Chen Jing, Ai Yong, Ye Demao, et al. Signal fading under 16km horizontal laser link aperture reception conditions [J]. Semiconductor optoelectronics, 2008, 29(06): 928–931.
10. Prokes, Ales. Atmospheric effects on availability of free space optics systems [J]. Optical Engineering, 2009, 48(6): 635–644.
11. Chen Chunyi, Yang Huamin, Jiang Huilin, et al. Analysis of large-aperture reception performance and aperture size selection in atmospheric optical communication [J]. China Laser, 2009, 36(11): 2957–2961.
12. Lou Yan, Jiang Huilin, Chen Chunyi, et al. Experimental analysis of light intensity fluctuation and spot area for laser atmospheric turbulence transmission [J]. Infrared and Laser Engineering, 2011, 40(03): 515–519.
13. LU Hongqiang, WANG Lahu, LI Yang, et al. Computational simulation of aperture averaging effect in space laser communication [J]. Applied Optics, 2012, 33(03): 619–623.
14. Wang Zongxing. Research on large aperture and composite aperture receiving antennas in optical wireless communication [D]. Harbin: Harbin Institute of Technology, 2015.
15. Li Zhipeng, Yang Guowei, Bi Meihua. Study of spatial diversity to suppress atmospheric scintillation index of vortex beams [J]. Optical Communication Technology, 2017, 41(02): 48–51.

16. Soni G, Tripathi A, Mandloi A, et al. Compensating rain induced impairments in terrestrial FSO links using aperture averaging and receiver diversity [J]. Optical and Quantum Electronics, 2019, 51(7): 1–11.
17. Dai Tianjun. Long-range Atmospheric Laser Communication System and Atmospheric Turbulence Suppression Technology [D]. Changchun: Changchun University of Science and Technology, 2021.
18. Yang Changqi, Jiang Wenhan, Rao Changhui. Effect of aperture averaging on the BER of free-space optical communication [J]. Journal of Optics, 2007, 299(02): 212–218.
19. Xi Xiaodong, Wu Hanping. Design of near-surface UV communication transmitting optical system based on Cassegrain structure [J]. Applied Optics, 2017, 38(02): 205–209.
20. Wu J, Boardman A. Coherence length of a Gaussian-Schell beam and atmospheric turbulence [J]. Journal of Modern Optics, 1991, 38(7): 1355–63.
21. Ricklin J C, Davidson F M. Atmospheric optical communication with a Gaussian-Schell beam [J]. Journal of the Optical Society of America, 2003, 20(5): 856–866.
22. Dogariu A, Amarande S. Propagation of partially coherent beams: turbulence-induced degradation [J]. Optics Letters. 2003, 28(1):10–12.
23. K D BG D V. Spatially partially coherent beam parameter optimization for free space optical communications [J]. Optics express, 2010, 18 (20): 20746–58.
24. Drexler K, Roggemann M, Voelz D. Use of a partially coherent transmitter beam to improve the statistics of received power in a free-space optical communication system: theory and experimental results [J]. Optical Engineering, 2011, 50(2): 025002.1–025002.7.
25. Zhang Lei, Chen Ziyang, Xiong Mengsu, et al. Scintillation index of partially coherent light transmission in atmospheric turbulence [J]. Intense Lasers and Particle Beams, 2014, 26(09): 77–81.
26. Ke Xizheng, Zhang Ya, Chen Wei. Effect of non-Kolmogorov atmospheric turbulence on the Rayleigh interval and turbulence distance of radial partially coherent array beams [J]. Journal of Photonics, 2016, 45 (11): 23–30.
27. Zhang Bo. Laguerre-related beam modulation and its atmospheric turbulence transport characteristics [D]. Nanjing: Nanjing University of Technology, 2020.
28. Yu J, Zhu X, Wang F, et al. Experimental study of reducing beam wander bymodulating the coherence structure of structured light beams [J]. Optics Letters, 2019, 44(17): 4371–4374.
29. Zhang Jie. Partial coherent optical atmospheric turbulence transmission beam parameter simulation technique [D]. Changchun: Changchun University of Technology, 2019.
30. Yang Ning, Zhao Liang, Xu Ying, et al. Transmission characteristics of partially coherent radially polarized distorted beams in non-uniform atmospheric turbulence [J]. Laser and Infrared, 2022, 52(08):1167–1176.
31. Ke Xizheng, Deng Lijun. Theory of partially coherent optical transmission in optical wireless communication [M]. Beijing: Science Press, 2016.
32. Ke Xizheng, Han Meimiao, Wang Mingjun. Spreading and drifting of partially coherent light on horizontal transmission paths in atmospheric turbulence [J]. Journal of Optics, 2014, 34(11):78–82.
33. Ke Xizheng, Han Meimiao, Wang Mingjun. Spreading and drifting of partially coherent light on the oblique-range transport path in atmospheric turbulence [J]. Journal of Photonics, 2015, 44(03): 149–153.
34. Ke Xizheng, Zhang Yu. Light intensity scintillation effect of partially coherent light in atmospheric turbulence [J]. Journal of Optics, 2015, 35(01):56–62.
35. Ke Xizheng, Wang Wanting. Light intensity and expansion of partially coherent light in oblique range and horizontal atmospheric turbulence [J]. Journal of Applied Sciences, 2015, 33(02):142–154.
36. Ke Xizheng, Wang Wanting. Beam extension and angular expansion of partially coherent light in atmospheric turbulence [J]. Infrared and Laser Engineering, 2015, 44(09): 2726–2733.

37. Ke Xizheng, Wang Jiao. Comparison of upward and downward transmission polarization characteristics of partially coherent beams in atmospheric turbulence [J]. Journal of Physics, 2015, 64(22):148–155.
38. Ke Xizheng, Wang Chaozhen. Light intensity distribution of partially coherent vortex beams during transport in atmospheric turbulence [J]. Advances in Lasers and Optoelectronics, 2016, 53(11): 83–91.
39. Ke Xizheng, Zhang Ya. Study of transmission characteristics of partially coherent array beams in optical wireless communication systems [J]. Advances in Lasers and Optoelectronics, 2016, 53(10):38–49.
40. Ke Xizheng, Zhang Huanjie. Study on the spreading characteristics of partially coherent Gaussian pulses in atmospheric turbulence [J]. Advances in Lasers and Optoelectronics, 2016, 53(08):110–116.
41. Wang Jiao, Ke Xizheng. Scattering characteristics of partially coherent beams transmitted in atmospheric turbulence [J]. Infrared and Laser Program, 2017, 46(07):135–142.
42. Ke Xizheng, Xue Yao. Effect of atmospheric turbulence scale on partially coherent optical transmission properties [J]. Journal of Photonics, 2017, 46(01):35–42.
43. Ke Xizheng, Xue Yao. Arrival angle undulation of partially coherent light in atmospheric turbulence [J]. Journal of Photonics, 2016, 45(12):30–34.
44. Ke Xizheng, Wang Song. Light intensity evolution of partially coherent Airy beams in atmospheric turbulence [J]. Journal of Photonics, 2017, 46(07):7–15.
45. Ke Xizheng, Zhang Lin. Beam extension and drift of partially coherent Airy light in atmospheric turbulence [J]. Journal of Photonics, 2017, 46(01):43–49.
46. Ke Xizheng, Wang Chaozhen. Light intensity distribution of partially coherent off-axis vortex beams in atmospheric turbulence [J]. Journal of Optics, 2017, 37(01):44–50.
47. Wang Jiao, Ke Xizheng, Wang Mingjun. Influence of source parameters and atmospheric turbulence on the polarization properties of partially coherent electromagnetic vortex beams [J]. Applied Optics. 2019, 58(24): 6486–6494
48. Wang Jiao, Ke Xizheng, Wang Mingjun. Detecting the topological charge of partially coherent electromagnetic vortex beams through the orientation angle of polarization [J]. Journal of Quantitative Spectroscopy & Radiative Transfer, 2019, 22(8): 11–16.
49. Gbur G, Wolf E. Spreading of partially coherent beams in random media [J]. Journal of the Optical Society of America, 2002, 19(8): 1592–1598.
50. Wolf E, Collett E. Partially coherent sources which produce the same far-field intensity distribution as a laser [J]. Optics Communications, 1978, 25(3): 293–6.
51. Chen Y, Gu J, Wang F, et al. Self-splitting properties of a Hermite-Gaussian correlated Schell-model beam [J]. Physical Review A, 2015, 91(1): 013823.1–013823.12.
52. Mei Z and Korotkova O. Cosine-Gaussian Schell-model sources [J]. Optics Letters, 2013, 38(14): 2578–2580.
53. Shu J, Xu H, Zhou Z L, et al. Beam Wander of the Multi-Gaussian Schell-Model Beam in Anisotropic Turbulence [J]. Progress In Electromagnetics Research, 2019, 78(4): 185–192.
54. M. M. lbrahim, A. M. lbrahim. Performance analysis of optical receivers with space diversity reception [J]. IEE Proceedings – Communications, 1996, 143 (6): 369–372.
55. Pan Feng, Ma Jing, Tan Liying, et al. Scintillation characterization of multiple transmitters for ground-to-satellite laser communication [J]. Infrared Components And Their Applications, 2004, 56(40): 448–454.
56. Sugianto Trisno, Igor I. Smolyaninov, Stuart D. Milner, et al. Characterization of time delayed diversity to mitigate fading in atmospheric turbulence channels [J]. University of Maryland/ College Park (United States), 2005, 58(92): 589215-1–589215.10.
57. A. J A, A. M N, V. B V. Multi-beam space-time coded systems for optical atmospheric channels [J]. Free-Space Laser Communications, 2006, 63(04): 63041B.1–63041B.9.
58. Cvijetic Neda, Wilson Stephen G, Brandt-Pearce Maite. Performance bounds for free-space optical MIMO systems with APD receivers in atmospheric turbulence [J]. IEEE Journal on Selected Areas in Communications, 2008, 26 (3): 3–11.

59. Ke Xizheng, Song Peng, Pei Guoqiang. Research on multi-aperture reception technology in wireless laser communication [J]. Journal of Optics, 2011, 31 (12): 22–28.
60. Ke Xizheng, Chen Juan, Pei Guoqiang. Research on multi-beam emission technology in wireless laser communication [J]. Optical Engineering, 2012, 39(07):1–7.
61. Jiang Long. Research on the performance and application of spatial diversity technology in free-space optical communication [D]. Beijing: Beijing University of Posts and Telecommunications, 2014.
62. Kaur P, Jain V K, Kar S. Performance analysis of free space optical links using multi-input multi-output and aperture averaging in presence of turbulence and various weather conditions [J]. IET Communications, 2015, 9(8): 1104–1109.
63. Bhatnagar M R, Ghassemlooy Z. Performance analysis of gamma–gamma fading FSO MIMO links with pointing errors [J]. Journal of Lightwave technology, 2016, 34 (9): 2158–2169.
64. Meng Desai. Performance study of outdoor visible light communication in weak turbulent channels [D]. Xi'an: Chang'an University, 2017.
65. Kumar N, Khandelwal V. Simulation of MIMO-FSO System with Gamma-Gamma Fading under Different Atmospheric Turbulence Conditions [A]. IEEE. 2019 International Conference on Signal Processing and Communication [C]. IEEE, 2019: 117–124.
66. Xu Yiyun. Research on field-of-view range expansion technology for space laser communication reception [D]. Changchun: Changchun University of Technology, 2019.
67. Ding KeNan. Research on Diversity Technology in Atmospheric Partial Coherent Optical Communication [D]. Hubei: Huazhong University of Science and Technology, 2021.
68. Liu, Zheqi. Research on atmospheric transmission characteristics based on wireless optical MIMO [D]. Changchun: Changchun University of Technology, 2022.
69. Ke Xizheng, Liu Mei. Diversity reception technique in optical wireless communication for turbulent channels [J]. Journal of Optics, 2015, 35(01): 88–95.
70. R K Tyson. Adaptive optics and ground-to-space laser communications [J]. Applied Optics, 1996, 35 (19): 3640–3646.
71. Roggemann M C, Lee D J. Two deformable mirror concept for correcting scintillation effects in laser beam projection [J]. Michigan Technological University, 1998, 33(81): 66–75.
72. Pierre R Barbier, David W Rush, Mark L Plett, et al. Performance improvement of a laser communication link incorporating adaptive optics [J]. University of Maryland/College Park (United States), 1998, 34(32): 93–102.
73. Li Xinyang, Jiang Wenhan. Control performance analysis of two adaptive optical systems with tandem correction [J]. Journal of Optics, 2001, 21 (09): 1059–1064.
74. Weyrauch T, Vorontsov M A. Free-space laser communications with adaptive optics: atmospheric compensation experiments [J]. Journal of Optical and Fiber Communications Reports, 2004, 1(4): 355–379.
75. Xia Lijun, Li Xiaofeng. Wavefront correction experiments for atmospheric optical communication based on adaptive optics [J]. Information and Electronics Engineering, 2010, 8(03): 331–335.
76. Han Liqiang. Performance and compensation method of spatial optical communication under atmospheric turbulence [D]. Harbin: Harbin Institute of Technology, 2013.
77. Wu Yunyun. Research on the application of adaptive optics in atmospheric optical communication [D]. Beijing: Graduate School of Chinese Academy of Sciences (Institute of Optoelectronics Technology), 2013.
78. Hashmi A J, Eftekhar A A, Adibi A, et al. Analysis of adaptive optics-based telescope arrays in a deep-space inter-planetary optical communications link between Earth and Mars [J]. Optics Communications, 2014, 33(3): 120–128.
79. Cvijetic M, Li M. Coherent free space optics communications over the maritime atmosphere with use of adaptive optics for beam wavefront correction [J]. Applied Optics, 2015, 54 (6): 1453–1462.

80. Wright M W, Morris J F, Kovalik J M, et al. Adaptive optics correction into single mode fiber for a low earth orbiting space to ground optical communication link using the OPALS downlink [J]. Optics Express, 2015, 23(26): 33705–33712.
81. Escarate P, Aguero J C, Zuniga S, et al. Model Predictive Control for Laser Beam Shaping [J]. IEEE Latin America Transactions, 2017, 15(4):626–631.
82. Li Beibei. Research on turbulence suppression technology in spatial optical communication [D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
83. Lin, Haiqi. Research on the control technology of adaptive optical system based on model recognition [D]. Chengdu: University of Chinese Academy of Sciences (Institute of Optoelectronics Technology, Chinese Academy of Sciences), 2019.
84. Chen Mo. Research on key technologies of ground station for satellite-terrestrial laser communication based on large-aperture telescope [D]. Chengdu: University of Chinese Academy of Sciences (Institute of Optoelectronics Technology, Chinese Academy of Sciences), 2019.
85. Banet Matthias T and Spencer Mark F. Compensated-beacon adaptive optics using least-squares phase reconstruction [J]. Optics express, 2020, 28(24): 36902–36914.
86. Yang, Huizhen, Su, Hang, Zhang, Zhi. Wavefront correction of SPGD control algorithm based on K-L mode [J/OL]. China Laser: 1–14 [2023-05-10].
87. Zhaokun L, Tao S, Xiongchao L, et al. Advanced multi-feedback stochastic parallel gradient descent wavefront correction in free-space optical communication [J]. Optics Communications, 2023, 53(3):43–51.
88. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress on adaptive optics for optical wireless communication systems at Xi'an University of Technology [J]. Intense Lasers and Particle Beams, 2021, 33 (08): 30–52.
89. Wu, Jaili, Ke, Xizheng. Adaptive optical correction of wavefront-free sensors [J]. Advances in Lasers and Optoelectronics, 2018, 55(03): 030103-1–030103-7.
90. Wang Xia Yao. Research on adaptive optical correction technology for vortex beams [D]. Xi'an: Xi'an University of Technology, 2018.
91. Ke Xizheng, Wang Xiaoyao. Experimental study on wavefront aberration correction of vortex light [J]. Journal of Optics, 2018, 38(03): 197–203.
92. Ke X Z, Zhang D. Fuzzy control algorithm for adaptive optical system [J]. Applied Optics, 2019, 58(36): 9967–9975.
93. Ke Xizheng, Yang Ke, Zhang Ying. Experimental study of wavefront distortion detection by CAPIS technique [J]. Advances in Lasers and Optoelectronics, 2019, 56(12): 30–37.
94. Yang Ke. Research on wavefront sensing technology for computational optical field imaging [D]. Xi'an: Xi'an University of Technology, 2019.
95. Ke Xizheng, Zhang Yunfeng, Zhang Ying, et al. Graphics processor acceleration for wavefront-free sensing adaptive wavefront correction systems [J]. Advances in Lasers and Optoelectronics, 2019, 56(07): 96–104.
96. Ke Xizheng, Han Ke Na. Wavefront simulation and wavefront correction for liquid crystal spatial light modulators [J]. Advances in Lasers and Optoelectronics, 2019, 56(05): 170–175.
97. Li Mei. Experimental study on the correction of beam front aberrations by the eigenmode method [D]. Xi'an: Xi'an University of Technology, 2020.
98. Cui Na-Mei. Experimental study on the correction of wavefront aberrations of vortex beams by phase difference method [D]. Xi'an: Xi'an University of Technology, 2020.
99. Cheng Shuang. Correction Techniques for Large Distortion Wavefront in Free Space Coherent Optical Communication [D]. Xi'an University of Technology, 2021.
100. Zhou RZ. Adaptive optics theory [M]. Beijing: Beijing Institute of Technology Press, 1996.
101. Wu Jiali. Experimental study of coherent optical communication system without wavefront detection [D]. Xi'an: Xi'an University of Technology, 2018.
102. Liu Jian. Characterization of transverse shear interference and wavefront reconstruction [D]. Liaoning: Dalian University of Technology, 2009.
103. Liu, Changjiang. Application simulation study of 16QAM coherent optical communication transmission system [D]. Beijing: Beijing University of Posts and Telecommunications, 2017.

104. Xu Qifeng. Adaptive optical control technology for wavefront curvature detection [D]. Hebei: North China University of Technology, 2020.
105. Medecki H, Tejnil E, Goldberg K A, et al. Phase-shifting point diffraction interferometer [J]. Optics Letters, 1996, 21 (19): 1526–1528.
106. Yao K N, Wang J L, Liu X Y, et al. Closed-loop adaptive optics system with a single liquid crystal spatial light modulator [J]. Optics Express, 2014, 22 (14): 17216–17226.
107. Liang JY, Wang HR, Zhang Na, et al. Research progress on deformation mirror and its control algorithm [J/OL]. Optical Communication Research: 1–17 [2023-04-27].
108. Zhang Danyu. Adaptive optical wavefront distortion control and experimental research [D]. Xi'an: Xi'an University of Technology, 2020.
109. Noll R J. Zernike polynomials and atmospheric turbulence [J]. Journal of the Optical Society of America, 1976, 66(03): 207–211.
110. Bei Rui. Adaptive optics with wavefront-free sensors for FSO systems [D]. Wuhan: Huazhong University of Science and Technology, 2011.
111. Zommer S, Ribak E N, Lipson S G, et al. Simulated Annealing in Ocular Adaptive Optics [J]. Optics Letters, 2006, 31(7): 939–941.
112. Yang, Huizhen, Li, Xinyang, Jiang, Wenhan. Comparison of several stochastic parallel optimal control algorithms for adaptive optical systems [J]. Intense Lasers and Particle Beams, 2008, 20 (01):11–16.
113. Ke Xizheng. Space-time coding theory in wireless optical MIMO systems [M]. Beijing: Science Press, 2014.

Propagation of Partially Coherent Beams in Atmospheric Turbulence

16

Contents

Introduction 518
Research History and Current Situation 518
Foreign Research Progress 518
Domestic Research Progress 523
Gaussian–Schell Model of a Partially Coherent Beam 528
Properties of Partially Coherent Beam Propagation in Atmospheric Turbulence 529
Intensity Distribution 529
Beam Spreading 533
Beam Wandering 537
GSM Beam Scintillation 539
Angle-of-Arrival Fluctuation 541
Polarization Characteristics 544
Speckle Characteristics 548
Summary 550
References 552

Abstract

When light waves propagate in the atmosphere, due to the random changes of atmosphere refractive index, turbulence effects such as scintillation, beam expansion, and beam wander will be caused, which seriously affect the performance of laser communication and tracking and imaging systems. Partially coherent beam is less affected by atmospheric turbulence, which makes its propagation characteristics become a research hotspot in recent years. In this chapter, the research progress of partially coherent beam in atmospheric turbulence at home and abroad is summarized; the research work of Xi'an University of Technology in this field is also introduced, including the intensity distribution, beam spread, beam wander, scintillation, arrival angle fluctuation, polarization and scattering properties, etc. Finally, the development direction of the field is prospected.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_16

Keywords

Optical wireless communications · Partially coherent beam · Atmospheric turbulence · Propagating characteristics · Research progress

Introduction

Owing to the unique directionality, monochromaticity, and high coherence of lasers, laser technology has been increasingly used in the fields of laser guidance, tracking, laser weapons, laser communications, radar, laser ranging, aerial surveying and mapping, and satellite remote sensing and monitoring [1–7].

When a laser propagates in atmosphere, owing to the random fluctuation of the atmospheric refractive index, the beam may expand or drift, the angle of arrival may fluctuate, the light intensity may flicker, and other turbulence effects may occur. These seriously affect the stability and reliability of an atmospheric laser communication system [8–11]. The more coherent the laser is, the more it is affected by atmospheric turbulence; it is easily scattered, making the light intensity distribution uneven. Therefore, to improve the performance of optical communication, the turbulence effect must be reduced.

Partially coherent light, with low coherence, can better suppress these turbulence effects, which means that it should be investigated, theoretically and experimentally. In addition, the polarization state of partially coherent light changes when propagating in atmospheric turbulence. This change is related to atmospheric turbulence and the coherence of the light source [12].

In practical applications, with a long transmission distance and strong atmospheric turbulence, the coherence of laser light in atmospheric turbulence is severely damaged and it becomes partially coherent. Thus, the study of the transmission characteristics of partially coherent light in atmospheric turbulence is becoming increasingly important.

Research History and Current Situation

Foreign Research Progress

The theory of the transmission and scattering of optical waves in atmospheric turbulence can be traced back to the 1960s.

In 1960, Chernov [13] used a geometrical optics approximation method to construct the wave equation of a random continuous medium; however, the finite size of the wave beam was not considered. Therefore, the method had a limited application scope.

In 1961, Tatarskii [14] introduced the turbulence theory developed by Kolmogorov and Obukhov, and used the Rytov approximation method to describe an electric field in an exponential form and expand it into a series. By solving the

exponential series, the process of solving amplitude and phase fluctuations was simplified, and the experimental phenomena of light waves propagating in turbulent media were successfully explained. Therefore, this approximation method has become a classic theory for studying the propagation properties of light in weak turbulence.

After the 1970s, researchers began to experiment with different wave sources, such as sound waves, radio waves, lasers, etc., and found many new experimental phenomena, such as the saturation effect of light wave undulations and the rise-and-fall phenomenon. Unfortunately, Tatarskii's classic theory was not suitable for strong turbulence.

To better explain these phenomena, research theories for strong fluctuation were proposed, such as the Markov approximation [15], generalized Huygens–Fresnel principle [16], path-integration method [17], heuristic theory [18, 19], thin-screen theory [20], etc. The main analytical methods used are illustrated in Fig. 16.1 [21]. However, effective treatment methods are lacking for the critical fluctuation of weak and strong transitions, and it is difficult to obtain a universal theoretical system and results.

In 1970, Whitman et al. [22] derived an expansion formula for laser propagation in a random medium, based on the scalar wave equation, and provided a specific analysis using the Kolmogorov spectrum. They concluded that the larger the transmission distance, the larger the spreading radius.

In 1971, Lutomirski and Yura [16] applied the generalized Huygens–Fresnel integral method to the propagation of plane waves with finite apertures in turbulent media. This method easily dealt with the average optical intensity properties, beam spreading, and coherence degradation of different modes of beams in turbulence; therefore, it is still widely used.

In 1972, Klyatskin [15] proposed a Markov approximation method for solving the statistical moment equation of an optical field. It treated the light propagation process as a Markov process and then derived the equations of each order moment. This method well handled the problem of light propagation in strong turbulence.

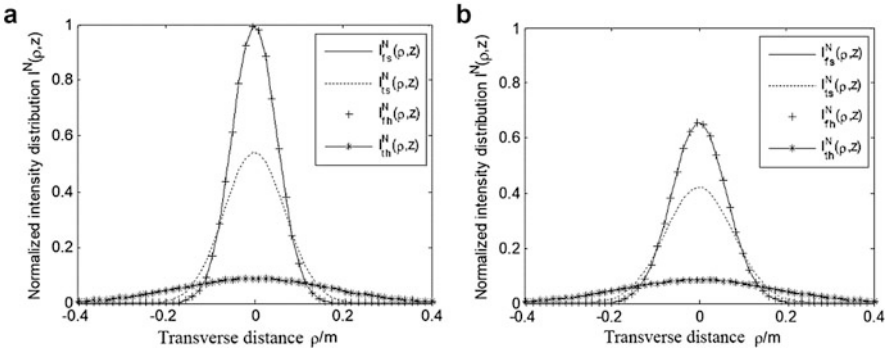


Fig. 16.1 Normalized intensity distribution of fully coherent and partially coherent GSM beams propagating in turbulence. (**a**) Fully coherent GSM beam. (**b**) Partially coherent GSM beam [101]

Poitier et al. [23] applied the Born approximation method to the spreading effect of a focused beam in turbulence.

In 1974, Clifford [24] proposed a simple physics-based, first-order perturbation method to investigate the beam scintillation saturation phenomenon under strong turbulence.

In 1975, Gardner et al. [25] studied the effects of turbulence on the propagation of pulsed laser beams. Fante [26] used the modified von Kármán spectrum to obtain a long-term beam expansion formula and derived a short-term beam expansion formula using the mutual coherence function. Richard J [27] used Ehrenfest's theorem to obtain an expression for the arrival angle of the beam in a turbulent medium with inhomogeneous weak undulation region and applied the Von Karman spectrum to analyze the arrival angle undulation drift of Gaussian focused and collimated beams.

In 1977, Wolf [28] proposed a typical partially coherent source, namely, the Gaussian–Schell mode beam, which has Gaussian intensity and spatial coherence distributions. Subsequently, Foley [29] studied the influence of the initial radius and coherence of the beam on its directionality.

In 1979, De Santis et al. [30] irradiated the light emitted by a laser onto rotating ground glass to generate pseudo-thermal light and formed a Gaussian–Schell mode beam through a Gaussian filter.

By the end of the 1970s, studies on the transmission properties of partially coherent beams in atmospheric turbulence had been initiated. The earliest was in 1978, when American scholar Leader [31] noticed the change in spatial coherence when he studied the scattering of laser irradiation on rough surfaces.

In 1983, Wang et al. [32] considered the intensity scintillation factor of a partially coherent beam transmitted through a weakly turbulent atmosphere. He smoothed the aperture at the receiving surface, using a method similar to that of Leader. Later, Soviet scholars Banakh et al. [33, 34] studied the scintillation of partially coherent beams in atmospheric turbulence.

In 1990, Wu et al. [35, 36] concluded that a partially coherent beam was less sensitive to turbulence than a fully coherent beam through the theoretical study of partially coherent Gaussian–Schell model (GSM) beam propagation in turbulent media. The conclusion greatly increased interest in partially coherent beams.

In 1999, Santarsiero [37] derived an expression for the quality factor of a Schell source beam. They analyzed the variation of the far-field intensity distribution and coherence of a partially coherent Gaussian–Schell beam and a J_0-correlated Schell-model beam during transmission.

Before 2000, research on the transmission properties of partially coherent beams in atmospheric turbulence was rare, and few results were published. With the rapid development of applications such as optical communication and light detection and ranging (LIDAR), the study of the transmission characteristics of partially coherent beams in atmospheric turbulence entered a new stage.

In 2002, to further verify Wu's conclusion, Ponomarenk et al. [38] and Gbur et al. [39] analyzed the propagation characteristics of partially coherent beams in

atmospheric turbulence, in terms of the radius and diffusion angle, respectively, and provided the standard and reason for this insensitivity. Subsequently, Ricklin et al. [40] used a partially coherent GSM beam as a laser source to study the properties of quasi-monochromatic partially coherent Gaussian beams during transmission, such as the average intensity, wavefront coherence length, and beam width, as well as the effects of the turbulence intensity and spatial coherence on an optical wireless communication link.

In 2003, Salem et al. [41] and Dogariu et al. [42] respectively proved theoretically and experimentally that partially coherent beams have more advantages for propagating in atmospheric turbulence. Shirai et al. [43] explained the conclusion that partially coherent beams are less affected by turbulence by studying the Hermi–Gaussian mode propagation of coherent Gaussian beams in atmosphere. Borghi et al. [44] discovered some coherence properties of partially coherent beams by measuring the intensity of partially coherent quasi-monochromatic light.

In 2004, Korotkova [45] studied the influence of a partially coherent light source and its signal-to-noise ratio (SNR) on the link SNR and bit error rate (BER), using theoretical analysis and experimental verification, and proved that the transmission performance of partially coherent light sources was better than that of coherent light sources under the same transmission channel.

In 2005, Andrews [46, 47] produced a drift model suitable for weak turbulence conditions, based on the Rytov and geometric optical approximations, introduced the Gaussian filter function, and attempted to extend it to a strong turbulence mode.

In 2006, Xiao [48] verified by modeling that partially coherent beam propagation in a weak-turbulent atmospheric channel effectively reduced the flicker index and improved the link-transmission quality.

In 2007, Lu et al. [49] studied the variation in the coherence of partially coherent electromagnetic Gaussian–Schell model beams. They found that with fixed source parameters and a specific atmospheric turbulence model, the coherence tended to maximize after propagating a certain distance. The effective width of the spectral coherence was also coherent.

In 2008, Mottagh [50] analyzed the influence of atmospheric turbulence on the SNR and BER of optical wireless communication systems under different turbulence conditions.

In 2010, Borah et al. [51] analyzed the relationship between parameters, such as the beam coherence length, curvature radius of the wavefront, and beam width, and the performance of a communication system. They found that using a partially coherent beam was more beneficial for improving the communication performance when the initial beam width was larger, the transmission distance was shorter, and the beam wavelength was shorter.

In 2012, Voelz et al. [52] studied the spot drift of focused partially coherent beams in atmospheric turbulence, based on a geometric optical approximation, and discussed the average intensity distribution and scintillation index of traceable and untraceable beams. The results were verified by numerical simulations.

In 2014, Chen [53] quantitatively studied the influence of turbulence on the diffusion of porous partially coherent beams by testing the relative mean-square width and obtained a turbulence-independent transmission range.

In 2015, Öztan and Baykal [54] carried out analytical derivations and numerical calculations of the scintillation index of partially coherent annular and flat-topped array beams under strong turbulence conditions. They concluded that the array beams had less scintillation than non-array beams. In the same year, Efimov [55] experimentally demonstrated that the scintillation index of partially coherent beams was lower than that of fully coherent beams at a transmission distance of 1.5 km–6.5 km.

Since 2016, the propagation of partially coherent beams in atmospheric turbulence has been studied extensively, and the beam polarization properties have also attracted attention. In the early days, the coherent and polarization properties of beams were investigated independently of each other. It was not until 1994 that James [56] first suggested that the polarization characteristics of partially coherent beams are altered when they are transmitted in free space.

In 1998, Gori [57] proposed a beam coherence–polarization matrix and calculated the polarization characteristics of spatial points in the time domain, which confirmed a conclusion similar to that of James.

In 2003, Wolf [58, 59] proposed a unified coherence–polarization theory, which was highly significant for studies on beam polarization properties. Subsequently, they used Young's interference experiment to verify that the difference in the degree of beam polarization at the detection plane was related to the degree of beam coherence at the incident end. They concluded that the polarization and coherence characteristics of the electromagnetic beam changed during the transmission in free space, which again proved the unified coherence–polarization theory [60, 61].

In 2005, Korotkova et al. [62, 63] used the unified coherence–polarization theory to derive a formula for the degree and angle of beam polarization during propagation and analyzed the change in the beam polarization state in free space and atmospheric turbulence.

In 2007, Eyyuboglu et al. [64] analyzed the variation in the polarization degree of partially coherent beams in atmospheric turbulence with the transmission distance, for different source parameters using a coherence–polarization matrix.

In 2009, Zhao et al. [65] investigated the polarization shift keying modulation technique in free-space optical communication systems and found that the BER could be reduced if the polarization state did not change during the beam transmission.

In 2014, Zhang [66] and Gao et al. [67] investigated the polarization properties of partially coherent Hermite–Gaussian beams propagating in a slanted path and partially coherent electromagnetic Gaussian–Schell pulsed beams in atmospheric turbulence, respectively. In 2019, they discussed the variation of polarization states of partially polarized and partially coherent pulsed electromagnetic beams as they propagate along a horizontal path in anisotropic turbulence. They compared and analyzed the change in spectral polarization between non-polarized, partially

polarized, and partially coherent pulsed electromagnetic Gaussian–Schell model beams in turbulence [68].

In 2021, Moreno and Castañeda [69] proposed the modeling of nonparaxial electromagnetic wave fields with partially spatially correlated amplitudes and polarization. They explained electromagnetic wave fields from the mathematics and physics perspectives (Table 16.1).

Domestic Research Progress

Many domestic scholars have extensively researched the transmission of partially coherent light in atmospheric turbulence.

In 1986, Zhang [70] discussed the influence of the spatial coherence degree of the light source, turbulence spectrum, and wavelength on the beam wander, and analyzed the drift variation of different turbulence intensities and transmission distances.

In 2004, Liu and Wu [71] derived and numerically simulated the scintillation index of a partially coherent beam under strong turbulence.

Feng et al. [72] studied the evolution of the polarization of a partially coherent beam propagating in random media, according to the generalized Huygens–Fresnel principle. They analyzed the influence of the coherence properties on the beam's polarization evolution and found that the polarization degree of partially coherent beams in random media was relatively less affected by random disturbances.

In 2005, Zhang and Tao [73] first derived a formula for the fluctuation variance of the angle of arrival of partially coherent Gaussian–Schell beams in atmospheric turbulence, and then derived an expression for the intensity distribution of a finite beam through atmospheric turbulence on a slanted path. The results showed that the average intensity was not only related to the half plane of the emission aperture, the transmission distance, and the long-term lateral coherence length but also to the radius of the emission aperture [74].

In 2006, Shi et al. [75] expressed the polarization state of a partially coherent electromagnetic beam propagating in a turbulent medium and analyzed the effect of turbulence on the polarization state of the beam.

In 2007, Gao et al. [76] developed a formula for the variance of beam wander suitable for a strongly fluctuating region based on that for weak fluctuation, which was in good agreement with the experimental data in the published literature. In addition, the spatial coherence length, wavelength, and turbulence intensity of the light source were numerically simulated for the wander, and the influence of these factors on the beam wander was discussed.

In the same year, Wang et al. [77] designed a spot drift detection system to detect the wandering of a He–Ne laser beam at close range. They adopted a charge-coupled device (CCD) camera with an image acquisition card to capture the output spot of the He–Ne laser at close range and used special software to process the digital image. Finally, the size of the spot drift was analyzed, as well

Table 16.1 Major methods of analyzing optical transmissions in atmospheric turbulence [21]

Theoretical research methods	Equation types	Processing procedure	Solving for parameters	Scope of application	Research applications
Rytov perturbation method	Parabolic equation	Assume that the optical field complex phase has a smooth change	Optic field distribution	Weak fluctuation, forward small-angle scattering	Amplitude and phase fluctuations
Geometrical optics method	Ray equation	Geometric optical approximation	Optic field distribution	Weak fluctuation, short-distance transmissions	Arrival angle fluctuation and beam wander
Huygens–Fresnel phase approximation	Generalized Huygens–Fresnel principle	Add a complex phase term to Green's function	Optic field distribution, optical field of moments of each order	Weak fluctuation, forward small-angle scattering	Finite-beam effect, imaging, phase correction analysis
Path-integration method	Parabolic equation	Solve parabolic equations by the path-integral method	Optic field distribution	Weak and strong fluctuation, forward small-angle scattering	Various effects, asymptotic analytical solutions
Feynman diagraming method	Green's function	Solve multiple-scattering equation using the Feynman diagram method	Optical field of moments of each order	Multi-scattering	Checking the reliability of various approximation methods
Statistical moments solved directly as Markov approximations	Statistical moment of the field equation	The view equation is a Markov process, assuming that only forward scattering exists	Moments of each order of the field	Strong fluctuation, forward small-angle scattering	Light intensity fluctuation, spatial distribution of light intensity

as the drift of the laser beam in the X and Y directions, and a spatial stereo angle was obtained.

In 2008, Qian et al. [78] numerically simulated the expansion and wandering of laser beams with different wavelengths using a multilayer phase screen method. The results showed that the larger the zenith angle, the stronger the turbulence effect, and the more the beam was affected by turbulence. The shorter the wavelength, the more the beam was affected by turbulence. The spot drift decreased with the increase of the initial beam radius and was independent of the wavelength.

Chu et al. [79] studied the average intensity and expansion of multi-Gaussian beams though atmospheric turbulence in a slanted path and analyzed the changes of intensity and expansion with parameters such as altitude, beam coherence length, and beam waist width. Wang et al. [80] studied the polarization characteristics of the entire optical field when an electromagnetic Gaussian–Schell model beam propagated in atmospheric turbulence.

In 2009, Yang et al. [81] investigated the beam width and angular expansions of a Hermite–Coulomb–Gaussian beam when it propagated horizontally in atmospheric turbulence. They analyzed the effects of the transmission distance, atmospheric refractive index structure constant, beam order, and beam waist width on the two expansions and provided corresponding physical explanations for the obtained results. Ji [82] studied the polarization, coherence, and spectral intensity of a partially coherent electromagnetic flat-topped beam propagating in atmospheric turbulence.

In 2010, Qian and Rao et al. [83] researched the beam expansion, wander, and scintillation of partially coherent beams by numerical simulation and once again verified that turbulence has less effect on partially coherent beams.

In 2011, Wu et al. [84] studied the propagation characteristics of electromagnetic partially coherent Gaussian–Schell beams in atmospheric turbulence with a slanted path. Chen et al. [85] analyzed and summarized the research progress of optical coherence and polarization.

In 2012, Zhang et al. [86] used a partially coherent flat-topped beam as an example to analyze the influence of parameters, such as the beam order, spatial coherence, and zenith angle, on the width of a beam that propagated in atmospheric turbulence on a slanted path. Li et al. [87] studied the expansion and wander of a GSM beam, using the beam center-of-mass distribution, and concluded that the longer the wavelength, the more obvious the wander.

In 2013, Duan et al. [88] studied the average intensity and mean-square beam width of a GSM beam when propagating in atmospheric turbulence with three different paths (uplink, downlink, and horizontal), and found that the beam was least affected by turbulence in the downlink path.

In 2014, Zhang et al. [89] studied the intensity distribution of a nonparaxial multicolor partially coherent Laguerre–Gaussian beam propagating in free space and found that the intensity distribution of the beam was determined by the initial parameters of the beam. They also found that the spectral displacement of the beam when propagating in free space was related to the initial beam parameters and the transmission distance.

In 2015, Fang et al. [90] experimentally generated a spatially coherent and controllable Besinc coherent vortex beam and studied the intensity and coherence distributions of the beam propagating in free space. Hou et al. [91] numerically calculated and analyzed the BER of a partially coherent optical transmission channel under different atmospheric turbulence conditions and proved that partially coherent optical transmissions effectively reduced the communication BER.

In 2017, Zhang et al. [92] studied the transmission characteristics of Gaussian inhomogeneous partially coherent beams in atmospheric turbulence and found that the change of its coherence was closely related to the Gaussian modulation half-width of the source plane and the transmission distance.

In 2018, Wang [93] summarized the research foundation and theoretical methods of partially coherent beams in atmospheric turbulence and introduced the phase screen numerical simulation method of a beam propagating through turbulent atmosphere.

In 2019, Zhou et al. [94] investigated the evolution rules of transmission characteristics, such as the intensity distribution, root-mean-square beam width, and beam wander, when a crescent-shaped beam was propagated in atmospheric turbulence.

In 2021, Zhang and Ma et al. [95] studied the effect of atmospheric turbulence on a twisted Laguerre–Gaussian correlated beam with different distortion factors, cross-sectional coherence, and beam orders.

In 2022, Yang et al. [96] studied the M2 factor of partially coherent twisted vortex beams in anisotropic atmospheric turbulence and conducted numerical simulations.

The transmission characteristics of partially coherent beams in atmospheric turbulence still require extensive research, and they are no longer limited to a single typical beam. In recent years, the following domestic universities and research institutes have investigated the transmission of partially coherent beams in atmospheric turbulence: University of Electronic Science and Technology of China, Changchun University of Technology, Xi'an University of Electronic Science and Technology, Taiyuan University of Technology, Anhui Institute of Optics and Precision Machinery, Institute of Optoelectronic Technology, Chinese Academy of Sciences, and Southwest Institute of Optoelectronics. The Optoelectronics Laboratory of Professor Ke Xizheng of Xi'an University of Technology has also studied the transmission characteristics of partially coherent beams in atmospheric turbulence and made progress [97].

In 2014, based on the formula of the beam width of a GSM beam in strong turbulence, Han [98] used the general expression of the beam wander variance models of Andrews and Phillips to derive the wander variance expression of a horizontal path propagating partially coherent beam in atmospheric turbulence. The spreading and wander of a partially coherent beam through the entire atmospheric turbulence were discussed.

In early 2015, based on the modified Rytov variance, Zhang [99] used a GSM beam and the Andrews model to study the effect of partially coherent beams through atmospheric turbulence on the log-intensity variance. In March of the same year, Han [100] further studied the spreading and wander of a coherent GSM beam while considering the outer turbulence scale in a slanted path.

Subsequently, Wang [101, 102] discussed the impact of the turbulence intensity, zenith, beam coherence length, and other parameters on a beam propagating on slanted and horizontal paths. They were compared from the perspective of the normalized intensity distribution and relative expansion. Wang [103, 104] compared and analyzed the influence of light source parameters and turbulence on the effective radius and average speckle radius of a GSM beam by numerically calculating their formulas.

In 2016, Xue [105, 106] studied the arrival angle fluctuation of partially coherent Gaussian–Schell beams in horizontal atmospheric turbulence. They contrasted and analyzed the influence of various parameters on the arrival angle fluctuation variance, including the inner and outer turbulence scales, turbulence intensity, transmission distance, source coherent parameter, and wavelength. Wang [107] analyzed and discussed the influence of the transmission distance, turbulence intensity, and beam parameters on the intensity distribution of a vortex beam.

Zhang [108] analyzed the transmission characteristics of partially coherent array beams in optical wireless communication systems. Zhang [109] conducted a detailed numerical analysis on the pulse spreading of partially coherent GSM pulsed beams propagating horizontally in atmospheric turbulence.

In 2017, Wang Song [110] simulated the effect of partially coherent Airy light intensity distribution in terms of propagation distance and turbulence intensity, and performed relevant experimental verification of the effect of the beam's own parameters on the light intensity distribution. Then, Lin Zhang [111] discussed and analyzed the expansion and drift characteristics of partially coherent Airy beams. Chaozhen Wang [112] extended the study of the transmission characteristics of off-axis vortex beams based on vortex beams.

In 2019, Wang [113, 114] studied the far-field distribution, intensity distribution, and polarization of a new partially coherent vortex beam after transmission through atmospheric turbulence. They also investigated the intensity distribution of a partially coherent vortex beam array transmitted in atmospheric turbulence.

During research, owing to factors such as the ground, high-altitude wind speeds, atmospheric temperature, and air pressure, the experimental conditions are limited, expensive, and difficult to achieve. Therefore, the transmission characteristics of partially coherent beams in atmospheric turbulence are mostly studied by theoretical analysis, numerical calculation, or software simulation. The Gaussian–Schell model (GSM), as a typical partially coherent beam, has excellent characteristics, such as more uniform light intensity and reduced

scattering; thus, it is typically studied when investigating partially coherent beam problems [115].

Gaussian–Schell Model of a Partially Coherent Beam

The coherence of a beam includes spatial and temporal coherence. Generally, a partially coherent light source is generated by superimposing a time-varying random amplitude and phase distribution onto a fully coherent beam field. The added random amplitude and random phase distribution represent the temporal and spatial coherence of the optical field.

Assuming that the optical field distribution of the initial fully coherent light source is D_{12}, the time-varying amplitude and phase distribution are expressed as $(\omega_1 - \omega_2)/(\omega_1 + \omega_2) \leq 1$. The optical field distribution of the partially coherent beam at the emission can be written as [117, 118]:

$$\begin{aligned} U'(x, y, 0, t) &= U_0(x, y, 0) t_A(x, y, t) \\ &= U_0(x, y, 0) \exp[i\xi(x, y, t)], \end{aligned} \tag{16.1}$$

where $\xi(x, y, t)$ is a random phase that varies with time and represents the partial coherence of the optical field of the light source.

The two-point cross-spectral density function of a partially coherent light source is defined as [40]:

$$\begin{aligned} W_0(\rho_1, \rho_2) = \left\langle U_0^*(\rho_1) U_0(\rho_2) \right\rangle &= U_0^*(\rho_1) U_0(\rho_2) \\ &< \exp[i\xi_1(\rho_1, t)] \exp[i\xi_2(\rho_2, t)] > , \end{aligned} \tag{16.2}$$

where $\rho_1 = \left(x_1^2 + y_1^2\right)^{1/2}$, $\rho_2 = \left(x_2^2 + y_2^2\right)^{1/2}$, and $\rho = \mid \rho_1 - \rho_2 \mid$ represent the two-dimensional vectors at the plane of the light source.

For partially coherent Schell beams, the cross-spectral density function is expressed as:

$$W_0(\rho_1, \rho_2) = \left\langle U_0^*(\rho_1) U_0(\rho_2) \right\rangle = \sqrt{I_0(\rho_1)} \times \sqrt{I_0(\rho_2)} \mu_0(\rho_2 - \rho_1), \tag{16.3}$$

where $I_0(\rho)$ is the average light intensity and $\mu_0(\rho_2 - \rho_1)$ represents the coherence degree of the beam at $z = 0$. When the beam satisfies

$$I_0(\rho) = A \exp\left(-\frac{2\rho^2}{w_0^2}\right), \mu_0(\rho) = \exp\left(-\frac{\rho^2}{2\sigma_g^2}\right), \tag{16.4}$$

it is called a partially coherent Gaussian–Schell model beam. A is a constant, w_0 is the initial beam radius, and σ_g represents the degree of beam coherence. The larger the σ_g, the higher the degree of coherence. In contrast, the coherence of the beam is worse.

Properties of Partially Coherent Beam Propagation in Atmospheric Turbulence

Intensity Distribution

When a beam is transmitted in atmospheric turbulence, owing to the random fluctuation of the atmospheric refractive index, the light intensity on the receiving surface weakens, which affects the reliability of the laser communication system.

Comparison of the Intensity Distributions of a GSM Beam on Horizontal and Slanted Transmission Paths

The transmission of a beam in atmospheric turbulence can be divided into horizontal and slanted paths, according to the size of the zenith angle. The beam is affected by turbulence to different degrees for different transmission paths. When the beam is propagating in free space, the light intensity is weakened, owing to the influence of vacuum diffraction, and the light intensity on the receiving surface will differ, depending on the beam coherence.

As shown in Fig. 16.1 [101], the normalized intensity distributions of fully coherent and partially coherent GSM beams in free space are independent of the transmission path, while the intensity distribution of a slanted path in atmospheric turbulence is more concentrated. Whether propagating in free space or in atmospheric turbulence, the intensity distribution of a fully coherent GSM beam is more concentrated. The intensity distribution of a partially coherent beam is less affected by turbulence than that of a fully coherent beam, which is consistent with the results obtained in the literature [43] when the beam is transmitted horizontally through the atmosphere.

Influence of Light Source Parameters on the Intensity Distribution

The normalized light intensity distribution during atmospheric turbulence propagation differs when the relevant parameters of the light source, such as beam coherence length and wavelength, are different. Figures 16.2, 16.3, and 16.4 [101] demonstrate the effects of various factors, such as the beam coherence length, beam waist width, and wavelength, on the normalized intensity distribution of partially coherent GSM beams propagating in atmospheric turbulence and free space.

It can be seen that when the GSM beam propagates in a slanted path, the smaller the beam coherence length, the smaller the influence of turbulence on the light intensity. The smaller the beam wavelength, the stronger its normalized light intensity, and the larger the wavelength, the stronger the light intensity. This is due to the different in-band attenuations caused by the atmospheric absorption of beams of different wavelengths.

Effect of Atmospheric Turbulence on the Intensity Distribution

The influence of atmospheric turbulence is generally studied theoretically by changing the atmospheric refractive index structure constant (atmospheric turbulence intensity) and the internal and external turbulence scales. Figures 16.5 and 16.6

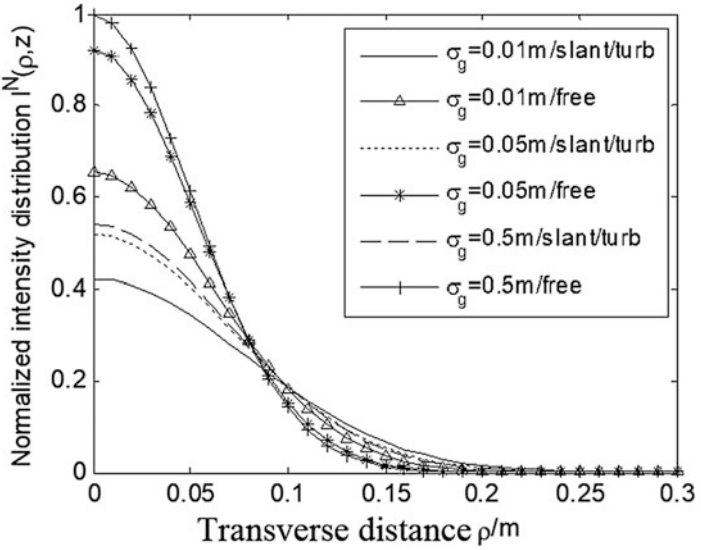


Fig. 16.2 Normalized intensity distribution of partially coherent GSM beam transmitted in turbulence with different coherent lengths [101]

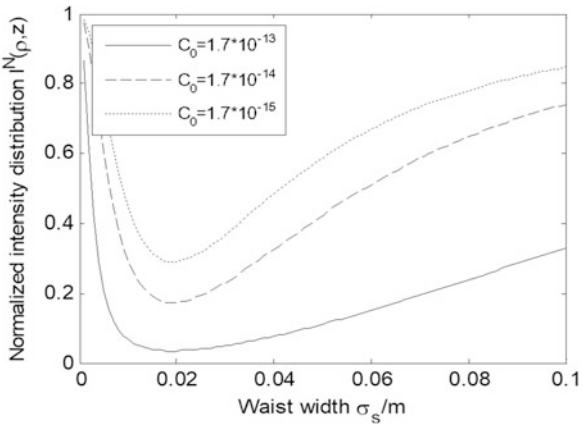


Fig. 16.3 Normalized intensity distribution of a partially coherent GSM beam transmitted in different turbulence conditions [101]

[101] show the normalized intensity distribution of partially coherent GSM beams propagating in atmospheric turbulence with different atmospheric structure constants and internal and external turbulence scales, respectively.

The results show that under the same turbulence condition, the intensity distribution of the beam on the slanted path is more concentrated than that of the horizontal path. The stronger the turbulence intensity, the more dispersed the intensity distribution; the larger the inner turbulence scale, the more concentrated

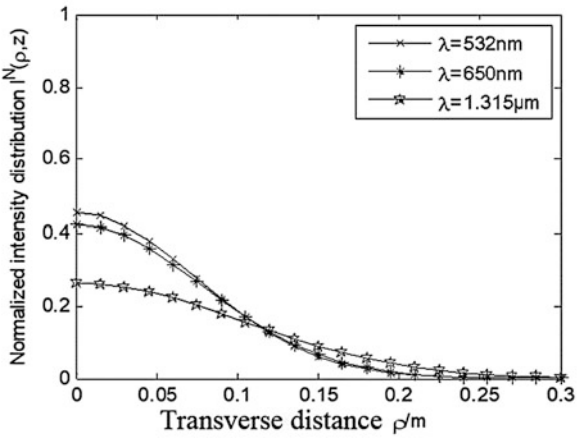


Fig. 16.4 Normalized intensity distribution of a partially coherent GSM beam transmitted through turbulence with different wavelengths [101]

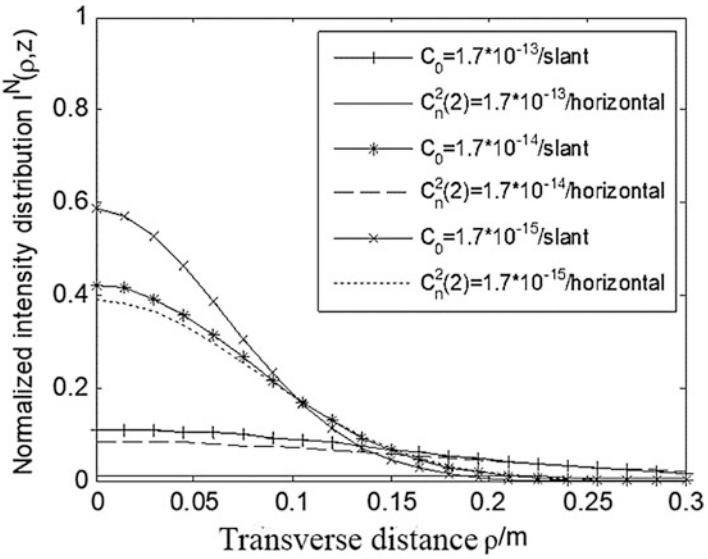


Fig. 16.5 Normalized light intensity distribution during partially coherent GSM beam propagation under different turbulence intensities [101]

the intensity distribution. Changes to the outer scale had little effect on the beam intensity distribution. Therefore, when analyzing the influence of atmospheric turbulence on the intensity distribution, only the internal scale need be considered; the outer scale can be ignored.

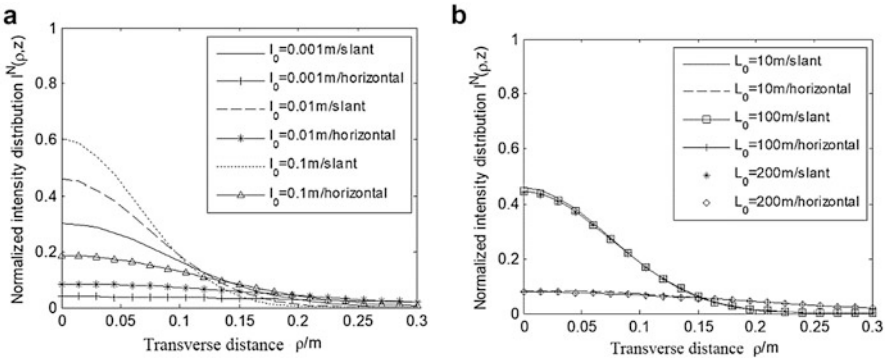


Fig. 16.6 Normalized intensity distribution of GSM beams transmitted in turbulence with different inner and outer scales. (**a**) Different inner scales. (**b**) Different outer scales [101]

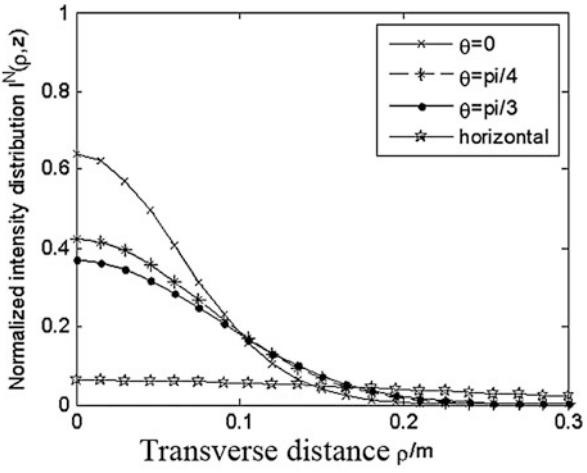


Fig. 16.7 Normalized intensity distribution of a partially coherent GSM beam propagating in turbulence with different zenith angles [101]

Effect of the Zenith Angle on the Intensity Distribution

When the beam passes through the atmosphere at different zenith angles, owing to different transmission distances, the influence of atmospheric attenuation and turbulence also differs. The normalized intensity distribution of a partially coherent GSM beam propagating in turbulence with different zenith angles is shown in Fig. 16.7 [101].

It can be seen from the figure that the zenith angle is 0°, i.e., when the beam propagates vertically, the intensity distribution is the most concentrated, while it is dispersed on the horizontal path. The larger the zenith angle, the more dispersed the light intensity distribution. The closer the beam is to the ground, the greater the influence of turbulence, and the worse the beam quality. Moreover, the larger the zenith angle, the more scattered the light intensity distribution.

Beam Spreading

Because of the influence of the atmospheric refractive index and beam wavefront phase distortion, the beam spreads when it propagates in atmospheric turbulence. The turbulence scale also affects it. When the beam scale is smaller than the turbulence scale, many small turbulence vortexes appear in the beam cross section. This part of the corresponding beam is diffracted, leading to random fluctuations of the beam in time and space, and eventually to spreading.

Spreading of a GSM Beam in Horizontal and Slanted Paths

The effective beam radius at the receiving end directly reflects the change in beam spreading. As shown in Figs. 16.8 and 16.9 [98] and Figs. 16.10 and 16.11 [100], when the beam is transmitted horizontally, the smaller the initial beam radius, the more the beam is affected by turbulence. A partially coherent beam transmitted horizontally is less affected by turbulence—i.e., it spreads less—than a fully coherent beam, which is consistent with the conclusions drawn in the literature [116], taking a partially coherent Hermite–Gaussian beam as an example.

During slanted propagation, a partially coherent beam expands more rapidly than a fully coherent beam and is also less affected by turbulence. The larger the initial radius and the higher the receiver height, the smaller the beam spreading, which is in agreement with the conclusions drawn from the numerical simulation of partially coherent beams in the literature [78].

In addition, Ke and Wang [102] discussed beam spreading from the perspective of the relative beam width (ratio of the beam width in turbulence to a free-space beam width) and concluded that when the beam propagates along a downlink, it is least affected by turbulence. The relative beam width increases with the increase in

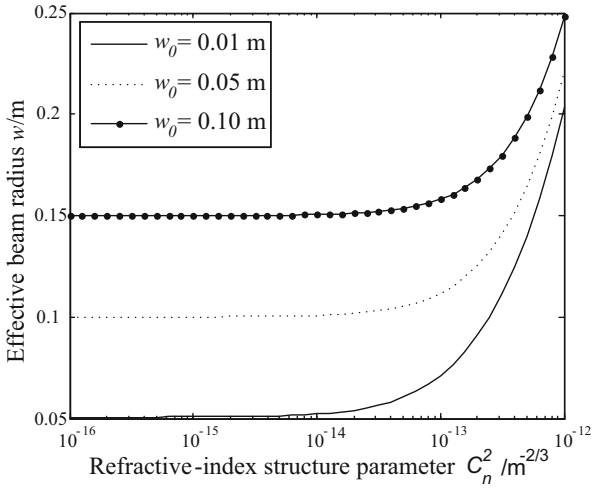


Fig. 16.8 Effective beam radius versus refractive index structure parameter [98]

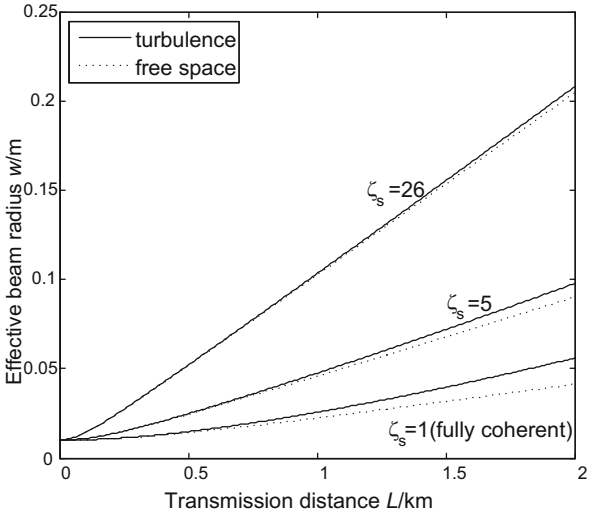


Fig. 16.9 Effective beam radius versus propagation distance [98]

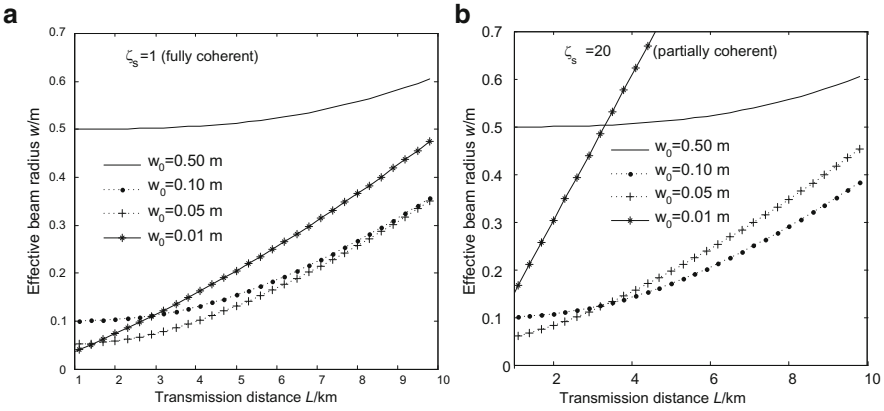


Fig. 16.10 Effective beam radius versus propagation distance with different initial beam radii. (**a**) Fully coherent beam. (**b**) Partially coherent beam [100]

turbulence intensity, zenith angle, and coherence length. It also increases with the decrease in wavelength and internal scale; thus, turbulence has a larger influence on the beam.

Angle Spreading of a GSM Beam in Horizontal and Slanted Paths

The angular spread variation can also reflect the spreading effect of atmospheric turbulence on the beam from another perspective. Figures 16.12, 16.13, 16.14, and 16.15 [102] demonstrate the effects of parameters, such as the zenith angle,

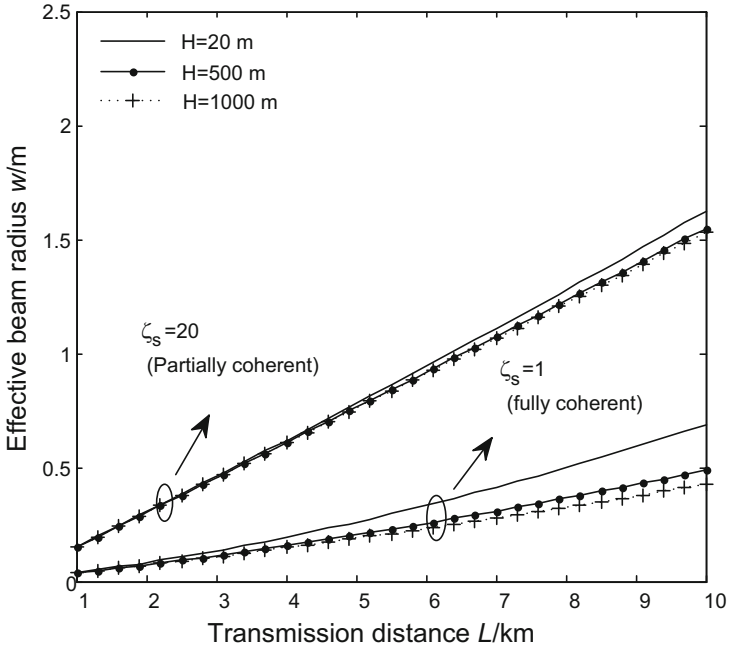


Fig. 16.11 Effective beam radius versus propagation distance with different receiver heights [100]

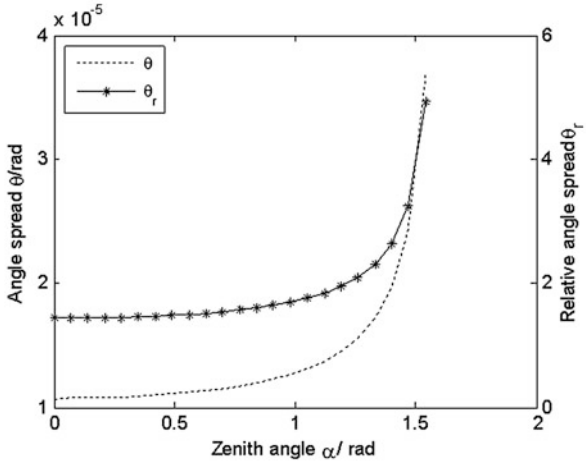


Fig. 16.12 Angular spread and relative angular spread versus zenith angle [102]

transmission distance, turbulence intensity, and coherence length, on the angular and relative angular spreading [119].

As can be seen from the figures, the angular spreading of the beam on the slanted path through atmospheric turbulence is less affected by the turbulence than that on

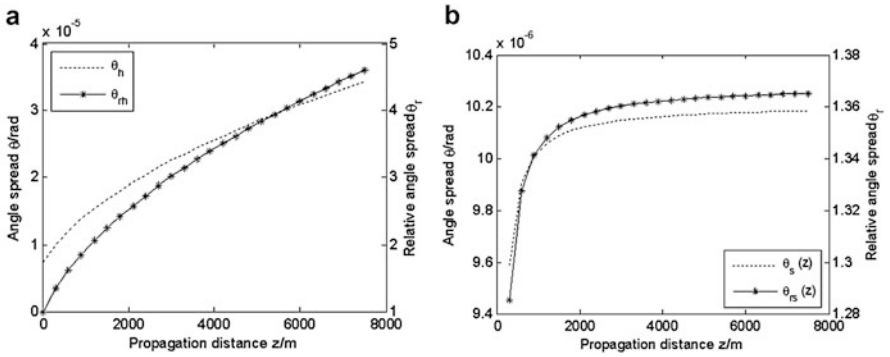


Fig. 16.13 Angular spread and relative angular spread versus propagation distance. (**a**) Horizontal transmission. (**b**) Slanted transmission [102]

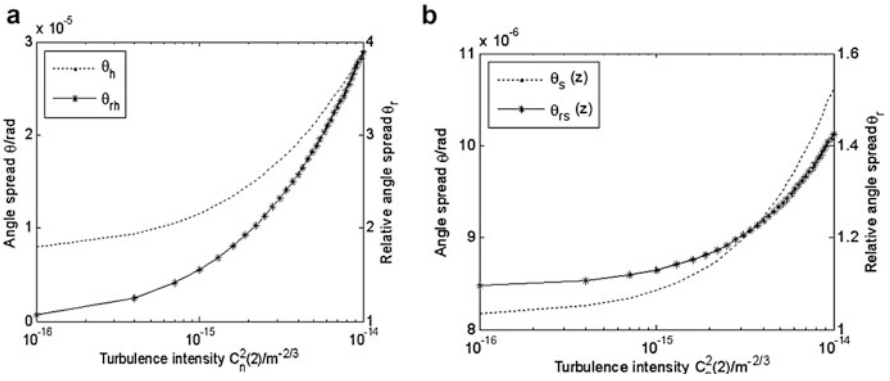


Fig. 16.14 Angular spread and relative angular spread versus turbulence intensity. (**a**) Horizontal transmission. (**b**) Slanted transmission [102]

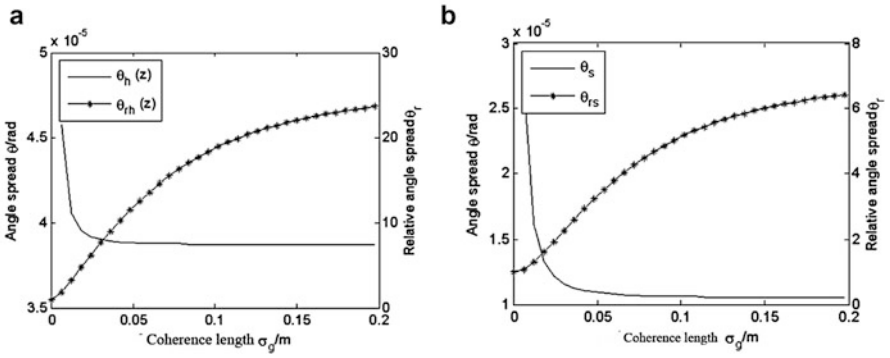


Fig. 16.15 Angular spread and relative angular spread versus coherence length. (**a**) Horizontal transmission. (**b**) Slanted transmission [102]

the horizontal path, and the smaller the zenith angle of the slanted transmission, the less the beam is affected by turbulence. The spreading of the beam angle and relative angle increase faster at transmission distances less than 1 km. The stronger the turbulence intensity, the greater the influence of turbulence on the beam angle spreading and the more obvious the angle spreading.

Beam Wandering

Beam wandering is mainly due to the refraction of large-scale turbulent vortices. When the beam diameter is much smaller than the turbulence size, i.e., when the turbulence is relatively weak, the turbulence causes random beam deflections and deviations from the original propagation direction. At this time, the beam propagation direction or the projection position on the receiving surface is randomly floating, i.e., the beam wanders.

Beam Wandering of a GSM Beam on a Horizontal Transmission Path

Beam wandering is inseparable from large-scale turbulence. A turbulence power spectral density function that considers large-scale factors needs to be selected in the research. As shown in Figs. 16.16 and 16.17 [98], the wavelength has little effect on wandering in the visible and near-infrared bands for short-range transmissions. The difference in wandering variation due to different wavelengths decreases as the initial radius increases. Therefore, to reduce the impact of atmospheric turbulence, the transmitter should choose a large initial radius, long wavelength, and higher position from the ground for transmitting.

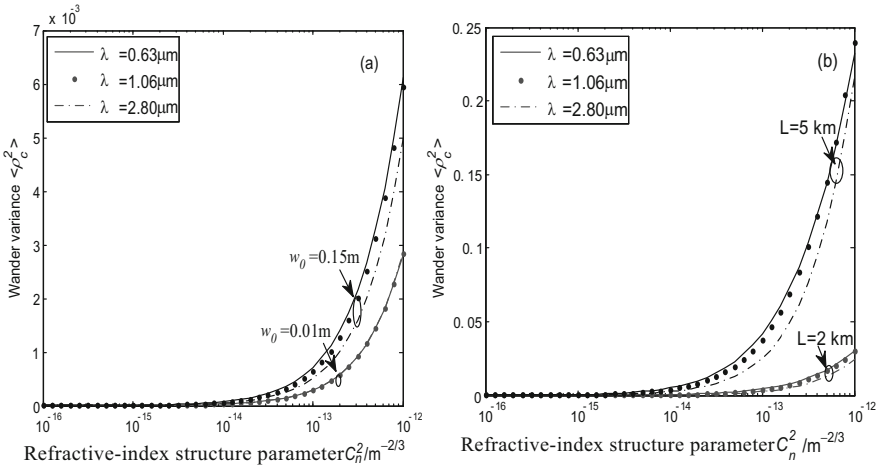


Fig. 16.16 Beam wandering variance versus refractive index structure parameter. (**a**) Different wavelengths and initial beam radii. (**b**) Different wavelengths and transmission distances [98]

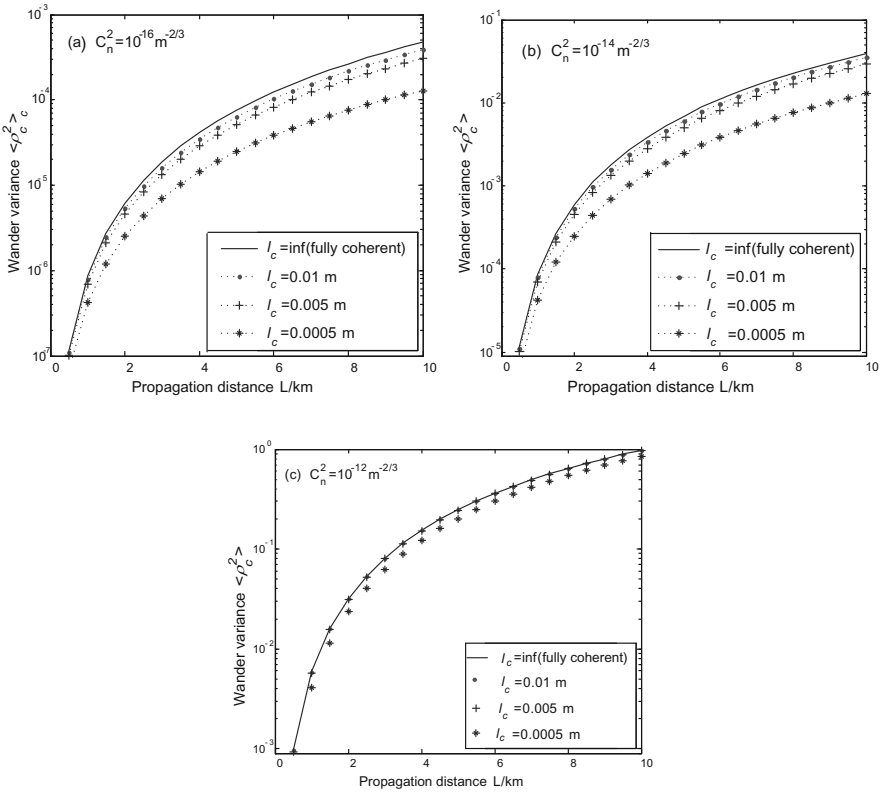


Fig. 16.17 Relationship of the beam wander variance and spatial coherence length. (**a**) $C_n^2 = 10^{-16}\mathrm{m}^{-2/3}$. (**b**) $C_n^2 = 10^{-14}\mathrm{m}^{-2/3}$. (**c**) $C_n^2 = 10^{-12}\mathrm{m}^{-2/3}$ [98]

When the transmission distance is greater than 2 km, the beam wandering in medium and strong turbulence is related to the wavelength, and the wandering is more obvious in the strong turbulence region. When the spatial coherence length of the beam is greater than 0.005 m, the spatial coherence length of the light source has little effect on the beam wandering within a transmission distance of 10 km, which is consistent with the conclusions in the literature [83].

GSM Beam Wandering on a Slanted Transmission Path

The atmospheric refractive index structure constant varies with height during a slanted transmission. Figure 16.18 [100] shows the variation in the wandering variance of partially coherent and fully coherent beams in atmospheric turbulence under different conditions during an uplink transmission. The beam wandering of the fully coherent beam is slightly affected by the wavelength, which is consistent with the numerical simulation results in the literature [83]. The longer the wavelength of the partially coherent beam, the more obvious the wandering, which agrees with the conclusion drawn from the spot centroid distribution diagram in the literature [87].

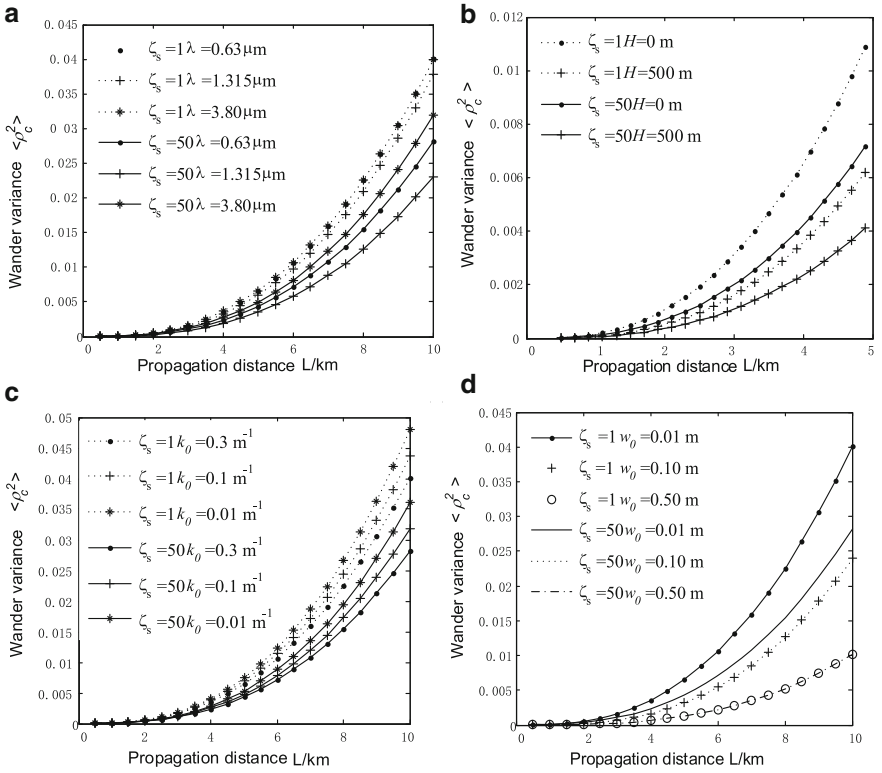


Fig. 16.18 Partially and fully coherent beam wander variance versus propagation distance. (**a**) Different wavelengths. (**b**) Different receiver heights. (**c**) Different large turbulence scales. (**d**) Different initial radii [100]

When the transmission distance is the same, the higher the height of the receiver, the smaller the wander effect.

Compared with a horizontal transmission, a slanted transmission is obviously less influenced by turbulence. When the receiver height is constant and the transmission distance is within 5 km, the worse the coherence, the smaller the wander variance, and the wandering is even smaller for short-distance transmissions. The larger the outer turbulence scale, the more apparent the wander is, and the partially coherent beam has a smaller wander effect than the fully coherent beam. When the transmission distance is greater than 2 km, the wander variance decreases with the increase of the initial beam radius.

GSM Beam Scintillation

When a laser propagates in atmospheric turbulence and the beam diameter is much larger than the turbulence scale, the beam cross section contains multiple turbulent

vortices, each of which independently scatters and diffracts the portion of the beam irradiated on it. This produces continuous spatial and temporal variations in the optical density in the detector plane. The beam intensity fluctuates from large to small, and the light spot is bright and dark. This change is called scintillation of the light intensity (light intensity fluctuation).

Scintillation in Strong and Weak Turbulence

According to Davis's classification approach, strong and weak turbulence are divided by the range of values of C_n^2: weak turbulence is when $C_n^2 < 6.4 \times 10^{-17} m^{-3/2}$ and strong turbulence is when $C_n^2 > 2.5 \times 10^{-13} m^{-3/2}$. The log-intensity variance of a partially coherent beam with the transmission distance under strong and weak turbulence is shown in Figs. 16.19 and 16.20 [99].

The log-intensity variance of a partially coherent beam is lower than that of a fully coherent beam under the same turbulence conditions. Moreover, as the light source coherence worsens, the log-intensity variance and the influence of the turbulence

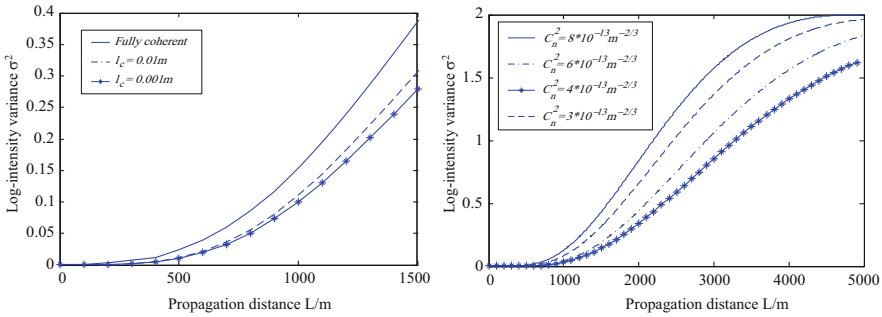


Fig. 16.19 Log-intensity variance of a partially coherent beam versus vector-propagation distance in strong turbulence [99]

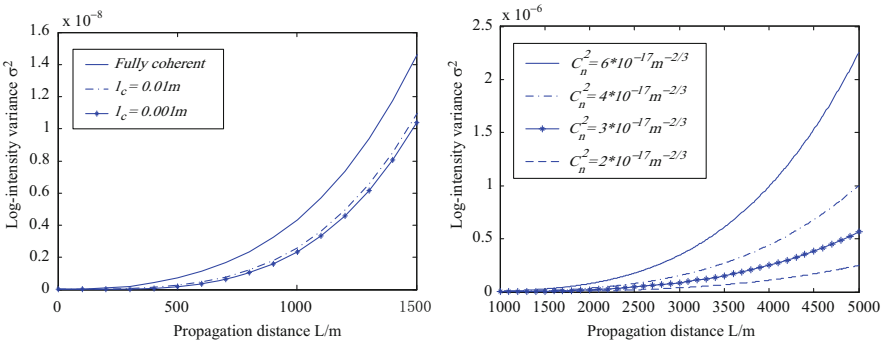


Fig. 16.20 Log-intensity variance of a partially coherent beam versus propagation distance in weak turbulence [99]

plane decrease. The light source coherence, atmospheric refractive index structure constant, and other parameters have a significant effect on the scintillation in both weak and strong turbulence. Regardless of turbulence, the scintillation of fully and partially coherent beams increases with the increase in transmission distance and atmospheric refractive index structure constant, and the partially coherent beam increases less than the fully coherent beam.

Scintillation of Fully and Partially Coherent Beams

To better compare and analyze the scintillation of partially and fully coherent beams in strong and weak turbulence, Ke and Zhang [99] introduced the modified log-intensity variance ratio α, which is defined as the ratio of the log-intensity variance of partially and fully coherent beams under the same conditions.

It can be observed from Fig. 16.21 that the value of α in weak turbulence is 0.91116 at a 2.5-km transmission distance, i.e., the log-intensity variance of a partially coherent beam is reduced by about 9.9%, compared with that of a fully coherent beam. The value of α in strong turbulence is 0.75281, i.e., the log-intensity variance of a partially coherent beam is reduced by about 24.7%, compared with that of a fully coherent beam. It can be seen that the partially coherent beam is significantly less affected by turbulence than the fully coherent.

Angle-of-Arrival Fluctuation

When a partially coherent Gaussian–Schell beam propagates in atmosphere, large and small vortices appear in its cross section, resulting in random disturbances that cause a phase shift in the wavefront. The phase change causes the angle of arrival to fluctuate. In practical applications, the angle-of-arrival

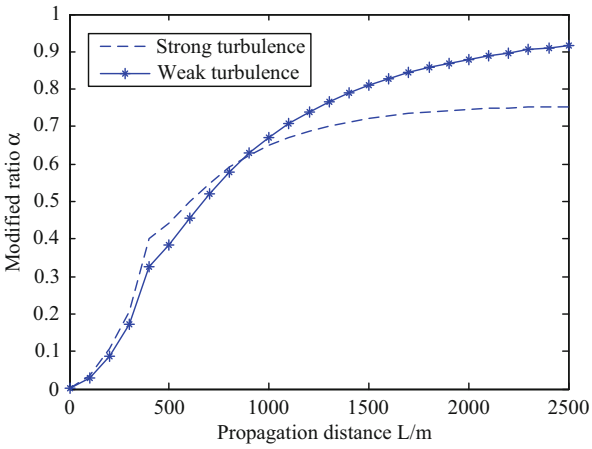


Fig. 16.21 Modified log-intensity variance ratio versus propagation distance [99]

fluctuation reduces the efficiency of coherent detection and can render a coherent-detection system invalid.

Effect of Light Source Parameters on the Angle-of-Arrival Fluctuation

The angle-of-arrival fluctuation of a partially coherent beam transmitted along a horizontal path in atmospheric turbulence differs depending on the coherent parameters and wavelengths of the light source. This is shown in Figs. 16.22 and 16.23 [106], respectively. The results show that the greater the wavelength, the smaller the angle-of-arrival fluctuation.

The angle-of-arrival fluctuation of a partially coherent Gaussian–Schell beam in atmospheric turbulence reaches a maximum at a certain initial radius. When the

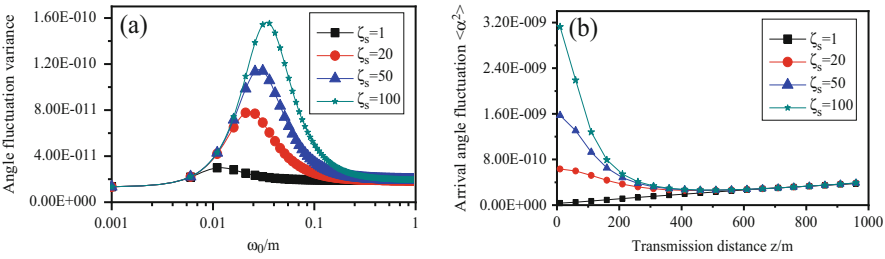


Fig. 16.22 Angle-of-arrival fluctuation with (**a**) waist radius and (**b**) transmission distance under different source coherence parameters [106]

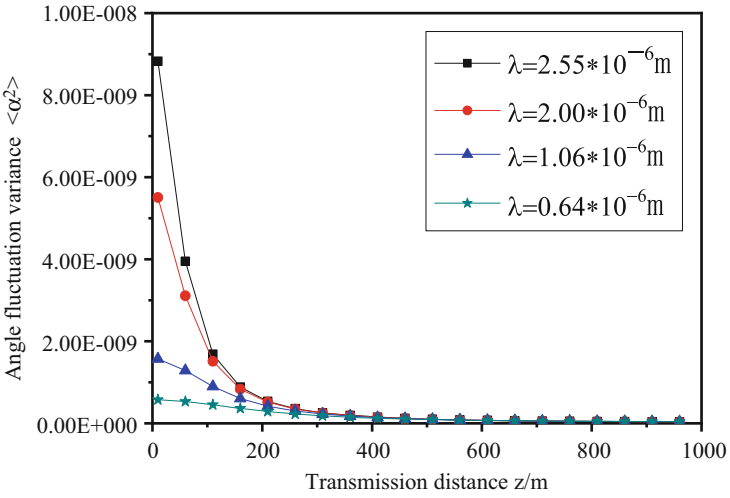


Fig. 16.23 Angle-of-arrival fluctuation variance with the transmission distance at different wavelengths [106]

initial radius of the beam waist is greater than 0.01 m, the arrival angle fluctuation of the beam varies significantly with the initial radius; the larger the beam coherence, the more obvious the phenomenon. As the transmission distance increases, the arrival angle fluctuation increasingly decreases. Compared to a partially coherent beam, the angle-of-arrival fluctuation of a fully coherent beam is less affected by turbulence.

Effect of Atmospheric Turbulence on the Angle-of-Arrival Fluctuation

In addition to the influencing light source parameters, the turbulence intensity and inner and outer turbulence scales also affect the arrival angle fluctuation. From the simulation results in Figs. 16.24 and 16.25 [106], it can be concluded that when the transmission distance is constant, the beam angle-of-arrival fluctuation first increases and then decreases. The smaller the inner turbulence scale, the greater the impact on

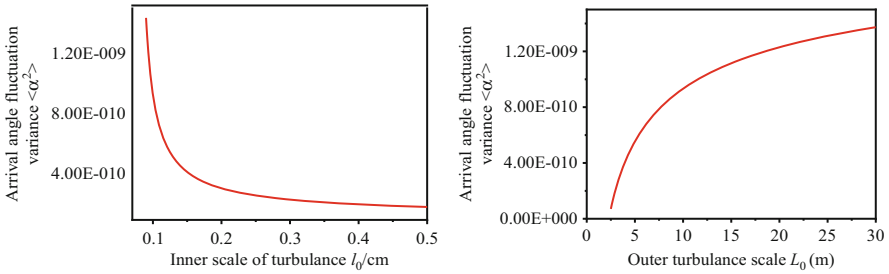


Fig. 16.24 Angle-of-arrival fluctuation with different internal and external scales [106]

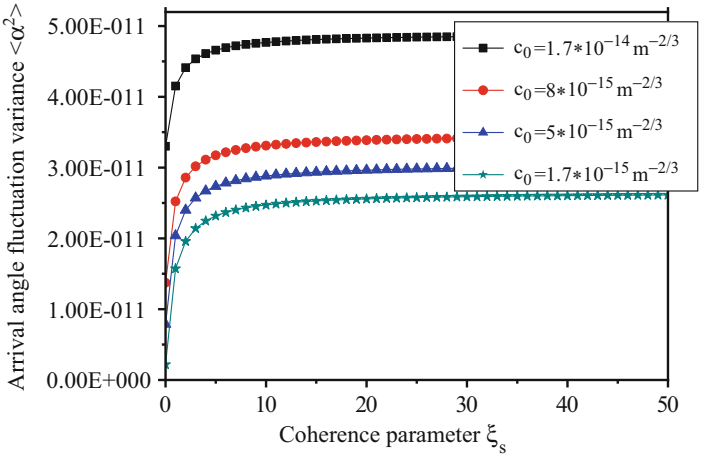


Fig. 16.25 Angle-of-arrival fluctuation variance with different source coherence parameters and atmospheric refractive index structure constants [106]

the angle-of-arrival fluctuation; the larger the outer scale, the greater the angle-of-arrival fluctuation. When the source coherence parameter is greater than 10, the variance of the angle-of-arrival fluctuation changes very little. The greater the turbulence intensity, the greater the angle-of-arrival fluctuation.

Polarization Characteristics

Existing studies have shown that the polarization characteristics of a beam in atmospheric turbulence change to various extents [56], and such variation is closely related to the coherence of the light source component and turbulence intensity. Polarization control is indispensable in wireless laser coherent optical communication systems. Only when the polarization states of the signal beam and local oscillator beam are the same or similar can the coherent light-detection sensitivity be high.

Polarization Characteristics of Partially Coherent EGSM Beams in Uplink and Downlink Paths

The degree of polarization (DoP) and polarization angle are two important parameters that describe the polarization characteristics; therefore, they are often used for the theoretical analysis of the polarization characteristics of various beams. Ke and Wang [104] analyzed the polarization characteristics in uplink and downlink paths by calculating the two parameters of a partially coherent electromagnetic Gaussian–Schell model (EGSM) beam propagating through turbulence.

Comparison of the Polarization Degree

The distribution curves of the DoP at the receiving end for partially coherent EGSM beams transmitted in uplink and downlink paths through atmospheric turbulence are shown in Figs. 16.26 and 16.27 [103]. This shows that, under

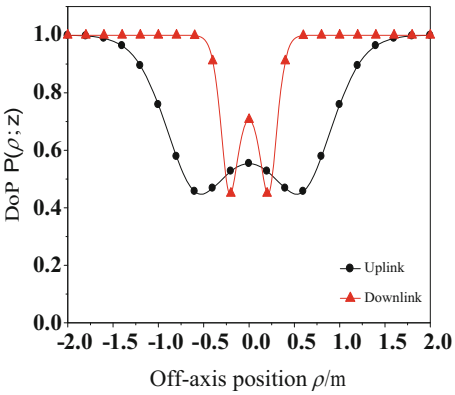


Fig. 16.26 Distribution curve of the DoP in uplink and downlink transmissions [103]

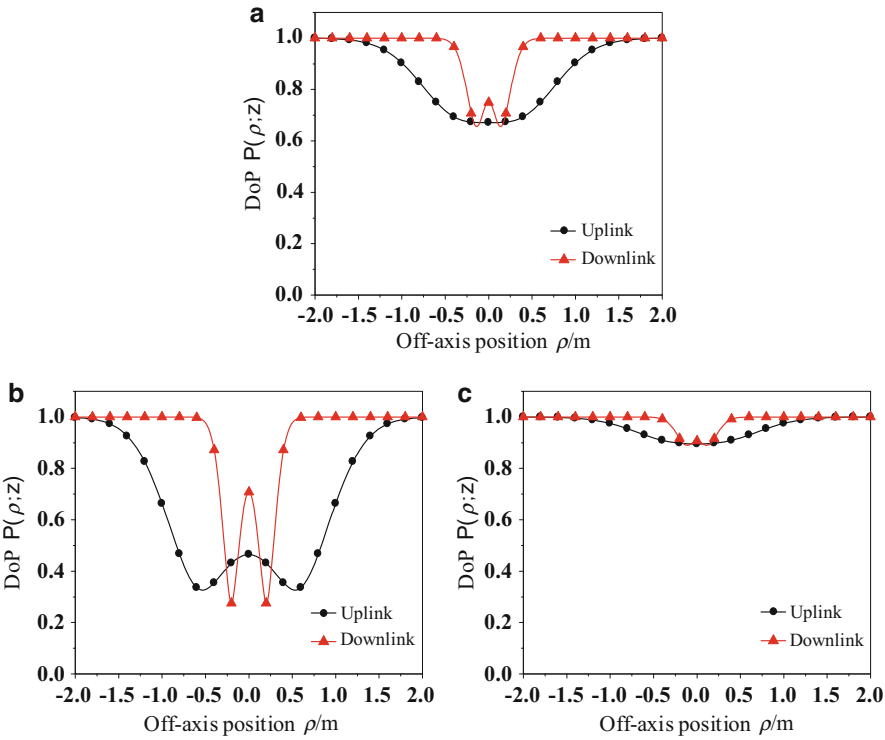


Fig. 16.27 Distribution curve of DoP in uplink and downlink transmissions when (**a**) $A_x/A_y = 1$, (**b**) $A_x/A_y = 2$, and (**c**) $A_x/A_y = 1/2$ [103]

the same conditions, the DoP distribution at the receiving end is more concentrated for downlink transmissions of an EGSM beam than for uplink transmissions.

At a transmission distance of 10 km, whether it is an uplink or downlink path, the DoP at the receiving end first decreases, with an increase in the deviation from the axial position, and then increases and approaches 1. This result is consistent with the conclusions of Fig. 3 in Ref. [6]. In addition, in Figs. 3 and 4 of Ref. [103], the DoP of the point on the receiving end axis is simulated and analyzed with different transmission distances. The results show that when the transmission distance is very short or very long, the DoPs of the uplink and downlink transmissions at the receiving end are almost the same for an equal amplitude ratio.

When the DoP of a point on the axis reaches the maximum value in a downlink transmission, the corresponding transmission distance is longer than the uplink transmission. In other words, when the EGSM beam is transmitted along the downlink path, the detector can receive the beam transmission information from

a longer distance. This is consistent with the results obtained in Fig. 2 in Ref. [62].

Comparison of Polarization Direction Angles

Figures 16.28 and 16.29 [103] show the polarization direction angle of a partially coherent EGSM beam transmitted in uplink and downlink paths to the receiving end with different transmission distances. It is obvious from the figure that in atmospheric turbulence, as the receiving height increases, the polarization direction angle in a downlink transmission changes faster than that in an uplink transmission. However, regardless of how it changes, the polarization direction angles of both uplink and downlink transmissions at the receiving end tend to have the same value at a long distance.

Polarization Comparison of Partially Coherent EGSM Beams on Horizontal and Slanted Paths

The parameters that affect the DoP distribution at the receiving end and the point change on the receiving end axis mainly include the initial beam radius, beam amplitude ratio, direction, wavelength, and turbulence intensity.

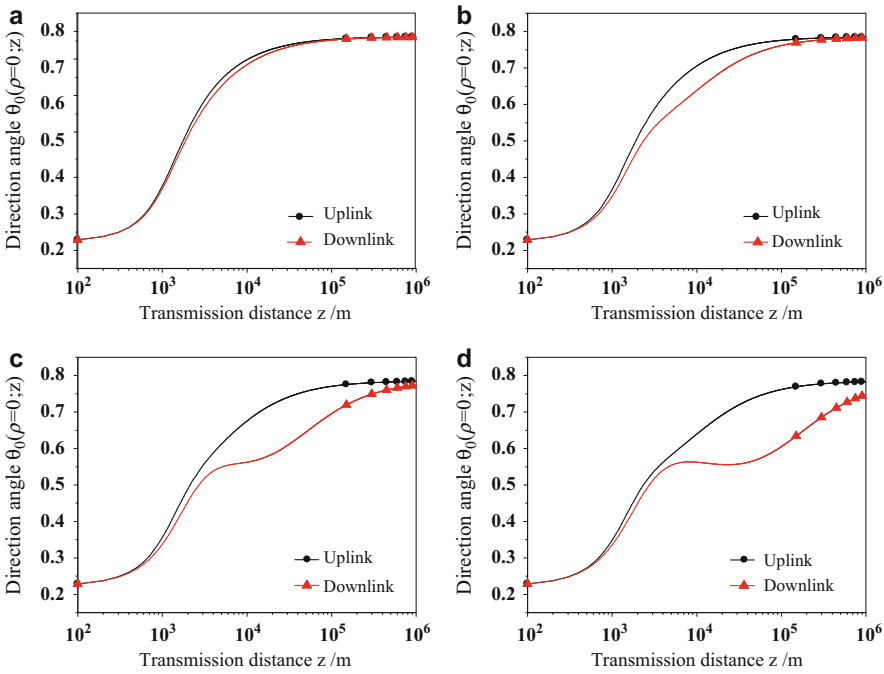


Fig. 16.28 Polarization direction angles with different transmission distances in uplink and downlink transmissions when $\sigma_x = \sigma_y = 0.01$ m. (**a**) H = 50 m; (**b**) H = 200 m; (**c**) H = 500 m; (**d**) H = 1000 m [103]

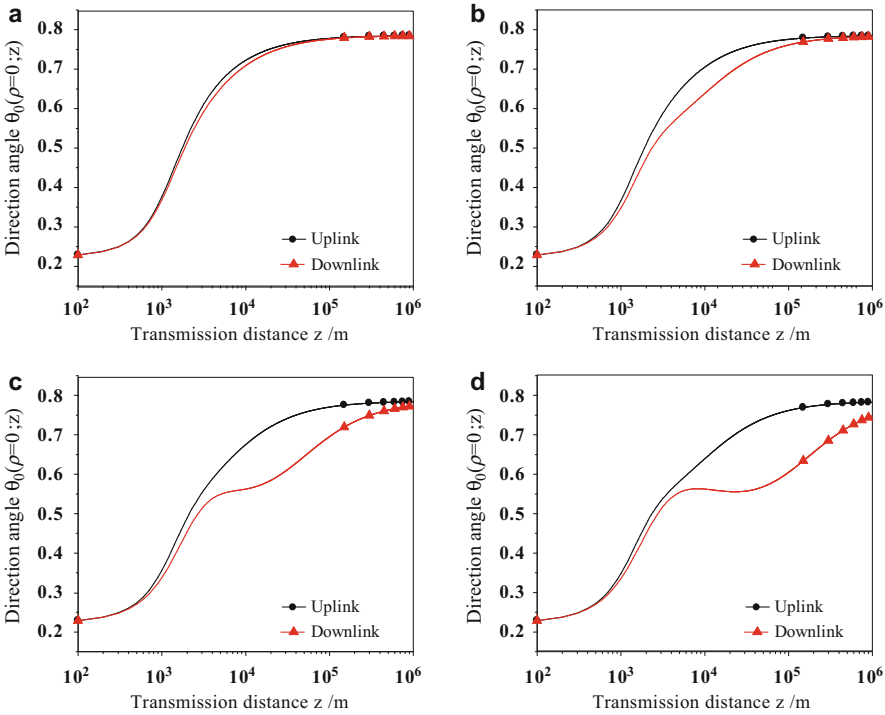


Fig. 16.29 Polarization direction angles with different transmission distances in uplink and downlink transmissions when $\sigma_x = 0.01$ m and $\sigma_y = 0.02$ m. (**a**) H = 50 m; (**b**) H = 200 m; (**c**) H = 500 m; (**d**) H = 1000 m [103]

1. Influence of the initial beam radius. When the initial radius components of the EGSM beam are equal, there is an optimal initial radius that most concentrates the distribution of the DoP at the receiving end. Regardless of how the initial beam radius changes, the DoP change of the entire optical field at the receiving end is more concentrated in a slanted transmission than in a horizontal transmission. Moreover, the corresponding transmission distance is longer when the DoP of the point on the receiving end axis reaches the maximum value. Whether the EGSM beam is transmitted along a slanted or horizontal path, the DoP at the receiving end first decreases with an increase in the distance from the axis, and then increases and tends to a fixed value of 1 [104].
2. Effect of the amplitude ratio. Regardless of how the amplitude ratio changes, the DoP change of the entire optical field at the receiving end in a slanted transmission is more concentrated than that in a horizontal transmission. When the DoP of the point on the receiving end axis reaches the maximum value, the corresponding transmission distance is longer than that in a horizontal transmission [104].

3. Effect of the turbulence intensity. Under the same turbulence condition, the DoP change of the entire optical field at the receiving end in a slanted transmission of an EGSM beam is more concentrated than that in a horizontal transmission. Regardless of a horizontal or slanted path, the DoP distribution of the entire optical field becomes more concentrated with the weakening of turbulence; however, the change in speed of a horizontal transmission is greater than that of a slanted transmission. As the transmission distance increases, the smaller the turbulence, the greater the variation of the DoP at the point on the axis [104].
4. Effect of the wavelength. When the beam wavelength is the same, the DoP change in the entire optical field at the receiving end in a slanted transmission of an EGSM beam is more concentrated than that in a horizontal transmission, and the detector can receive further information. As the wavelength increases, the DoP distribution of the entire optical field becomes more concentrated, regardless of the horizontal or slanted path. As the transmission distance increases, the longer the beam wavelength, the greater the variation of the DoP at the point on the axis [104].

Speckle Characteristics

When a beam propagates in atmospheric turbulence, the turbulence effect randomly modulates the beam intensity and phase and finally forms speckles in the far field. In space optical communication systems, the speckle effect directly affects the spot centroid detection accuracy of acquisition, pointing, and tracking (APT) systems. Therefore, it is necessary to conduct a detailed theoretical analysis of the speckle characteristics of the beam caused by atmospheric turbulence and provide a theoretical reference for the design of space optical communication systems.

Influence of Light Source Parameters

The average speckle radius of a beam is related to the initial radius, wavelength of the light source, and coherence length. Figures 16.30 and 16.31 [105] show the effects of the beam waist radius, coherence length, and light source wavelength on the effective beam radius and the average speckle radius, respectively. It is clear that the smaller the initial radius, the faster the beam spread, which is consistent with the conclusion given in reference [98].

Under the same transmission distance, the better the beam coherence, the larger the average speckle radius of the beam at the receiving end and the greater the beam diffusion. The smaller the wavelength, the smaller the effective beam radius and average speckle radius change with an increase in the transmission

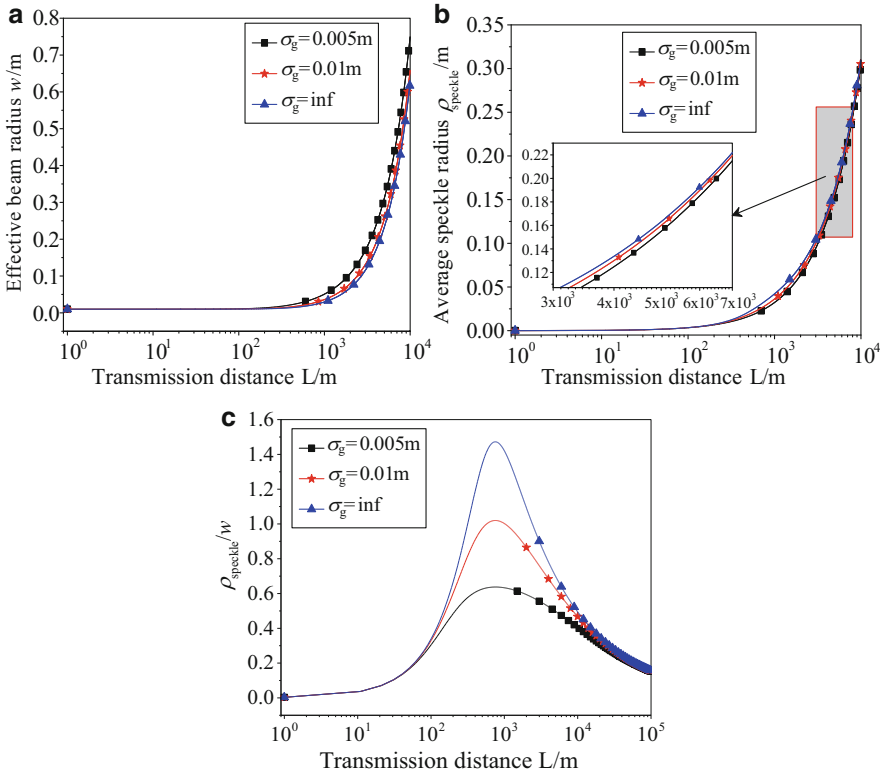


Fig. 16.30 Beam characteristics at the receiving end with different coherence lengths. (**a**) Effective beam radius; (**b**) average speckle radius; and (**c**) ratio of the average speckle radius to the effective beam radius [105]

distance, and the less the beam is affected when it propagates in atmospheric turbulence.

Influence of Atmospheric Turbulence Parameters

The turbulence intensity and inner and outer turbulence scales also affect the average speckle radius of the beam. Figures 16.32 and 16.33 [105] show that after the beam is transmitted for a certain distance, as the turbulence intensity increases, the beam spreads, owing to the phase distortion caused by the atmospheric refractive index fluctuation, which leads to a more dispersed spot at the receiving end. The effective beam radius and average speckle radius are greatly affected by the inner turbulence scale. The influence of the outer turbulence scale is negligible. This is consistent with the fact that the outer scale shown in Fig. 3 of Ref. [107] has little effect on the beam spreading.

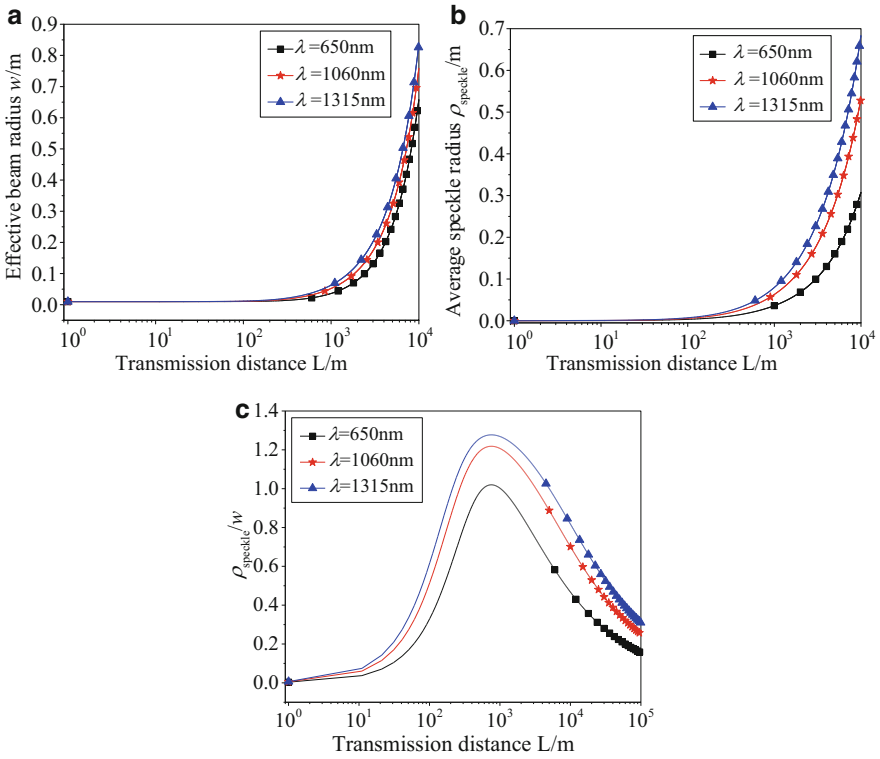


Fig. 16.31 Beam characteristics at the receiving end for different wavelengths. (**a**) Effective beam radius; (**b**) average speckle radius; and (**c**) ratio of the average speckle radius to the effective beam radius [105]

Summary

This chapter introduced the research progress of partially coherent beam propagation in atmospheric turbulence at home and abroad. It mainly focused on the research results of Xi'an University of Technology on the transmission characteristics of partially coherent GSM beams through the atmosphere in recent years. Compared with a fully coherent beam propagating in atmospheric turbulence, a partially coherent beam can suppress turbulence effects, such as beam spreading, light intensity fluctuation, and arrival angle fluctuations. Therefore, research on the effect of turbulence on partially coherent beams is of great significance.

In recent years, research on the transmission of partially coherent beams in the atmosphere has focused on the relationship between light source parameters and turbulence, partly to discuss and analyze the basic physical properties between different types of light sources and turbulence. More importantly, it has

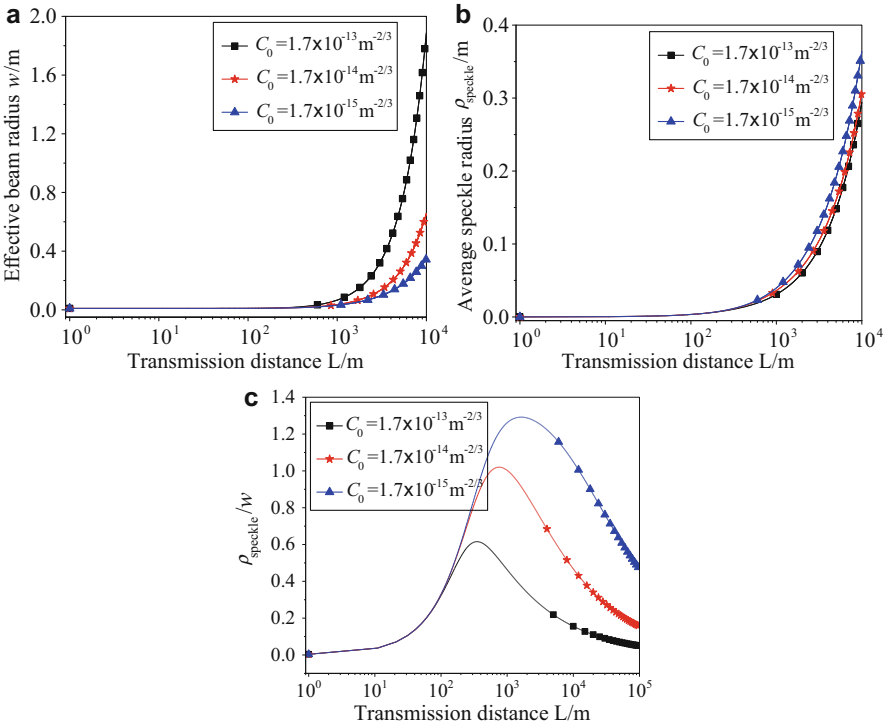


Fig. 16.32 Beam characteristics at the receiving end with different atmospheric refractive index structure constants. (**a**) Effective beam radius; (**b**) average speckle radius; and (**c**) ratio of the average speckle radius to the effective beam radius [105]

optimized the experimental results by changing the light source parameters in practical laser communication, such as improving the energy reception efficiency, reducing the light intensity scintillation, and reducing beam spreading and wandering.

However, most literature on partially coherent beam transmission properties is the result of theoretical calculations and simulation analyses, and most of them are ideal beams with definite mathematical analytical forms, such as partially coherent Gaussian–Schell beams, partially coherent flat-topped Gaussian beams, and partially coherent vortex beams. A large number of hypotheses and approximations have been used in the research process. The research results are still far from actual engineering applications and have certain limitations.

In addition, many of these findings have not yet been verified through numerical simulations or experiments. Although some verification methods have been proposed, they are generally inadequate. Therefore, whether theoretical, numerically simulated, or experimental, research on the transmission properties of partially coherent beams in atmospheric turbulence must be further improved.

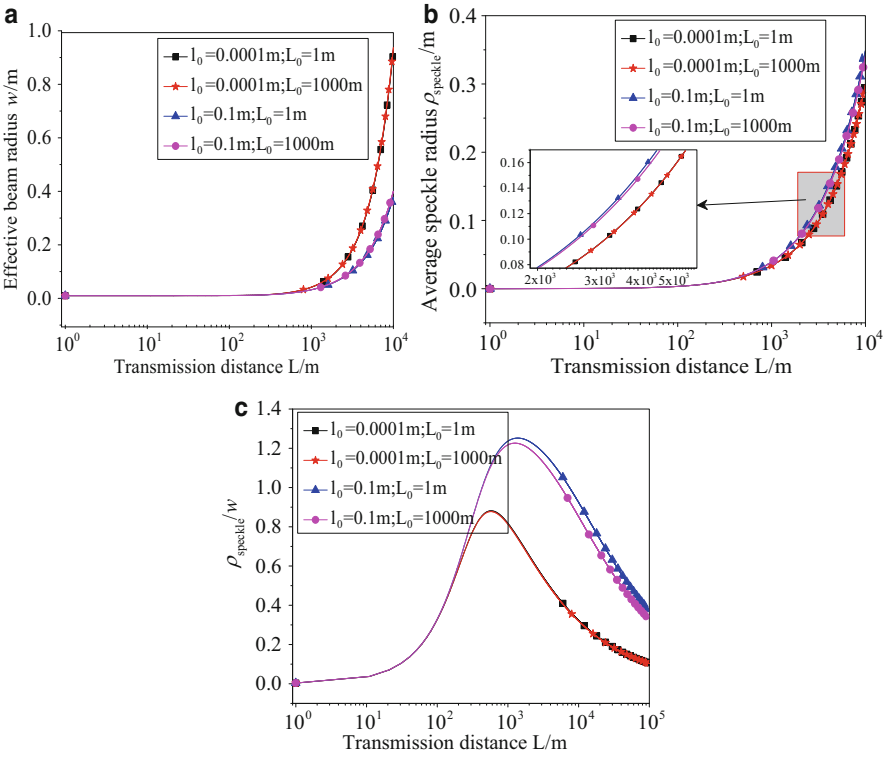


Fig. 16.33 Beam characteristics at the receiving end at different internal and external turbulence scales. (**a**) Effective beam radius; (**b**) average speckle radius; and (**c**) ratio of the average speckle radius to the effective beam radius [105]

References

1. Mandel L, Wolf E. Optical coherence and quantum optics [M]. New York: Cambridge University Press, 1995.
2. Zhang Yixin, Chi Zeying. Transmission and imaging of optical wave in atmosphere [M]. Beijing: National Defense Industry Press, 1998.
3. Wu Jian, Yang Chunping, Liu Jianbin. Theory of light transmission in atmosphere [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2005.
4. Ishimaru A. Wave Propagation and Scattering in Random Media [J]. Wave Propagation and Scattering in Random Media, 1978, 1(6):407–460.
5. Ke Xizheng, Xi Xiaoli. Introduction to wireless laser communication [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2004.
6. Cai Y, He S. Partially coherent flattened Gaussian beam and its paraxial propagation properties [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2006, 23(10): 2623–2629.
7. Nelson M, Avramov ZS, Korotkova O, et al. Scintillation reduction in pseudo Multi-Gaussian Schell-model beams in the maritime environment [J]. Optics Communications, 364 (2016): 145–1410.

8. Ke Xizheng, Yang Lihong, Ma Dongdong. Attenuation of laser signal transmission in rain [J]. Infrared and Laser Engineering, 2008, 37(06):1021–1024.
9. Yang Lihong, Ke Xizheng, Ma Dongdong, et al. Study on the transmission characteristics of 11.6 μm laser in sandstorm [J]. Journal of Xi'an University of Technology, 2009, 25(4): 472–476.
10. Yang Lihong, Ke Xizheng. Research on bit error rate based on polarization PPM of atmospheric optical communication [J]. Chinese Journal of Scientific Instrument, 2010,31(7): 1664–16610.
11. Rao Ruizhong. Propagation of light in turbulent atmosphere [M]. Hefei: Anhui Science and Technology Press, 2005.
12. Li Chengqiang, Wang Tingfeng, Zhang Heyong, et al. Effects of light source parameters and atmospheric turbulence on the polarization characteristics of electromagnetic beam propagation [J]. Acta Physica Sinica, 2014, 63(10):104201-1–104201-8
13. LA. Chernov. Wave propagation in a random medium [M]. New York: McGraw-Hill, 1960.
14. Tatarskii VI. Wave propagation in a turbulent medium [M]. New York: Dover Publication, 1961.
15. Klyatskin V I, Kon A I. On the displacement of spatially-bounded light beams in a turbulent medium in the Markovian-random-process approximation [J]. Radiophysics & Quantum Electronics, 1972, 15(9):1056–1061.
16. R F Lutomirski, et al. Wave Structure Function and Mutual Coherence Function of an Optical Wave in a Turbulent Atmosphere [J]. Journal of the Optical Society of America, 1971, 61(4): 482–488.
17. Wu Zhensen. Intensity correlation and scintillation of wave propagation saturation state in random medium [J]. Chinese Journal of Radio Science, 1986,01(01):76–89.
18. Hill R J. Theory of saturation of optical scintillation by strong turbulence - Plane-wave variance and covariance and spherical-wave covariance [J]. Journal of the Optical Society of America, 1982, 72(2):212–222.
19. Hill R J, Clifford S F. Theory of saturation of optical scintillation by strong turbulence for arbitrary refractive-index spectra [J]. Journal of the Optical Society of America (1917–1983), 1981, 71(6):675–686.
20. Flatte S M, Gerber J S. Irradiance-variance behavior by numerical simulation for plane-wave and spherical-wave optical propagation through strong turbulence [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2000, 17(6):1092–1098.
21. Wang Haiyan. Research on the transmission characteristics of partially coherent and partially polarized beam in atmospheric turbulence based on tensor method [D]. Nanjing University of Science and Technology, 2011.
22. Whitman A M, Beran M J. Beam Spread of Laser Light Propagating in a Random Medium* [J]. Journal of the Optical Society of America, 1970, 60(12): 1595–1602.
23. Poirier J L, Korff D. Beam Spreading in a Turbulent Medium [J]. Josa, 1972, 62(7):893–899.
24. Clifford S F, Ochs G R, Lawrence R S. Saturation of optical scintillation by strong turbulence [J]. J. opt. soc. am, 1974, 64(2):148–154.
25. Gardner C S, Plunus M A. The effects of atmospheric turbulence on the propagation of pulsed laser beams [J]. Radio Science, 1975, 10(1):129–137
26. Fante, R. L. Electromagnetic beam propagation in turbulent media [J]. Proceedings of the IEEE, 1975, 63(12): 1669–1692.
27. Cook R J. Beam wander in a turbulent medium: An application of Ehrenfest's theorem [J]. Journal of the Optical Society of America, 1975, 65(8):942–949.
28. Carter W H, Wolf E. Coherence and Radiometry with Quasi-Homogenous Planar Sources [J]. Journal of the Optical Society of America, 1977, 67(6):785–796.
29. Foley J. T., Zubairy M. S. The directionality of Gaussian-Schell model beams [J]. Optics Communications, 1978, 26(3):297–300.
30. Santis P D, Gori F, Guattari G, et al. An example of a Collett-Wolf source [J]. Optics Communications, 1979, 29(3): 256–260.

31. Leader J C. Atmospheric propagation of partially coherent radiation* [J]. Journal of the Optical Society of America, 1978, 68(2):175–185.
32. Baykal Y, Plonus MA, Wang S J. The scintillations for weak atmospheric turbulence using a partially coherent source [J]. Radio Science. 1983. 18(4):551–556.
33. Aksenov V P, Banakh V A, Buldakov V M, et al. Brief communications: Distribution of fluctuations of light intensity behind the objective of a telescope after reflection in a turbulent atmosphere [J]. Soviet Journal of Quantum Electronics, 1985, 15(10): 2136–1140.
34. Strong flux scintillations of incoherent radiation in a turbulent atmosphere [J]. Radiophysics & Quantum Electronics, 1981, 24(6):480–484.
35. J. Wu. Propagation of a Gaussian-Schell beam through turbulent media [J]. Journal of Modern Optics, 1990, 37(4): 671–684.
36. J. Wu and A. D. Boardman. Coherence length of a Gaussian-Schell model beam and atmospheric turbulence [J]. Journal of Modern Optics, 1991, 38(7): 1355–1363.
37. Santarsiero M, Gori F, Borghi R, et al. Spreading properties of beams radiated by partially coherent Schell-model sources [J]. Journal of the Optical Society of America A Optics & Image Science, 1999, 16(1):106–113.
38. Ponomarenko SA, Greffet J J, Wolf E. The diffusion of partially coherent beams in turbulent media [J]. Optics Communications, 2002, 208(1–3):1–9.
39. Gbur, Greg, Wolf, et al. Spreading of partially coherent beams in random media [J]. Journal of the Optical Society of America A, 2002, 19(8):1592–1599.
40. Atmospheric turbulence effects on a partially coherent Gaussian beam: implications for free-space laser communication [J]. J Opt Soc Am A Opt Image Sci Vis, 2002, 19(9):1794–1802.
41. Salem M, Shirai T, Dogariu A, et al. Long-distance propagation of partially coherent beams through atmospheric turbulence [J]. Optics Communications, 2003, 216 (4–6):261–265.
42. Dogariu A, Amarande S. Propagation of partially coherent beams: Turbulence-induced degradation [J]. Optics Letters, 2003, 28(1):10–13.
43. Shirai, Tomohiro, Dogariu, et al. Mode analysis of spreading of partially coherent beams propagating through atmospheric turbulence [J]. Journal of the Optical Society of America A, 2003, 20(6):1094–1102.
44. Borgi R, Guattari G, Torre L D L, et al. Evaluation of the spatial coherence of a light beam through transverse intensity measurements [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2003, 20(9):1763.
45. Korotkova, Olga. Model for a partially coherent Gaussian beam in atmospheric turbulence with application in Lasercom [J]. Optical Engineering, 2004, 43(2):330–341.
46. Andrews L C, Phillips R L. Laser Beam Propagation Through Random Media [M]. SPIE Press, 2005.
47. Andrews L C, Phillips R L, Sasiela R J, et al. Beam wander effects on the scintillation index of a focused beam [C], Proceedings of SPIE, 2005.
48. Xiao Xifeng,Voelz David. Wave optics simulation of pseudo-partially coherent beam propagation through turbulence: application to laser communications [J]. Free-Space Laser Communications VI,6304(2016): 63040L-1–63040L-7
49. Lu W, Liu L, Sun J, et al. Change in degree of coherence of partially coherent electromagnetic beams propagating through atmospheric turbulence [J]. Optics Communications, 2007, 271(1):1–9.
50. Motlagh A C, Ahmadi V, Ghassemlooy Z, et al. The effect of atmospheric turbulence on the performance of the free space optical communications [C], 2008 6th International Symposium on Communication Systems, Networks and Digital Signal Processing. IEEE, 2009.
51. Deva K. Borah, David G. Voelz. Spatially partially coherent beam parameter optimization for free space optical communications [J]. Optics Express, 2010, 20(18): 20746–20759.
52. Xiao X, Voelz D G. Beam wander analysis for focused partially coherent beams propagating in turbulence [J]. Optical Engineering, 2012, 51(2): 026001-1–026001-8
53. Xiaowen Chen,Mingyue Tang. Spreading of apertured partially coherent beams in turbulent media [J]. Optik, 2014,125(20),6032–6036.

54. Oztan, A Mehmet, Baykal, et al. Scintillations of partially coherent annular and flat-topped array beams in extremely strong turbulent medium [J]. Optics Communications, 354 (2015): 419–428.
55. Efimov A. Scintillations of a partially coherent beam in a laboratory turbulence: experiment and comparison to theory [J]. Proceedings of SPIE, 9354 (2015): 935404-1–935404-6.
56. James, Daniel F V. Change of polarization of light beams on propagation in free space [J]. Journal of the Optical Society of America A, 1994, 11(05):1641–1643.
57. F Gori, M Santarsiero, S Vicalvi, R Borghi, G Guattari. Beam coherence-polarization matrix [J]. Pure Appl. Opt, 1998, 7(05):941–951
58. Wolf E. Unified theory of coherence and polarization of random electromagnetic beams [J]. Physics Letters A, 2003, 312(05):263–267
59. E. Wolf. Correlation-induced changes in the degree of polarization, the degree of coherence, and the spectrum of random electromagnetic beams on propagation [J]. Optics Letters, 2003, 28(13): 1078–1080.
60. Shirai T, Wolf E. Coherence and polarization of electromagnetic beams modulated by random phase screens and their changes on propagation in free space [J]. Journal of the Optical Society of America A, 2004, 21(10): 1907–1916.
61. Roychowdhury H, Wolf E. Young's interference experiment with light of any state of coherence and of polarization [J]. Optics Communications, 252 (2005):268–274
62. Korotkova O, Wolf E. Changes in the state of polarization of a random electromagnetic beam on propagation [J]. Optics Communications, 246 (2005):35–43
63. Salem M, Korotkova O, Dogariu A, Wolf E. Polarization changes in partially coherent electromagnetic beams propagating through turbulent atmosphere [J]. Waves in Random and Complex Media, 2004, 14(4):513–523
64. Eyyubolu H T, Baykal Y, Cai Y. Degree of polarization for partially coherent general beams in turbulent atmosphere [J]. Applied Physics B, 2007, 89(1):91–98.
65. Zhao X, Yao Y, Sun Y, et al. Condition for Gaussian-Schell model beam to maintain the state of polarization on the propagation in free space [J]. Optics Express, 2009, 17(20):17888–17894.
66. Zhang L, Yin X, Zhu Y. Polarization fluctuations of partially coherent Hermite-Gaussian beams in a slant turbulent channel [J]. Optik, 2014, 125(13):3272–3276.
67. Ming G, Yan L, Hong L, et al. Polarization properties of polarized and partially coherent electromagnetic Gaussian-Schell model pulse beams on slant path in turbulent atmosphere [J]. Infrared Physics & Technology, 67 (2014): 98–106.
68. Li Y, Gao M. Turbulence-induced changes in the state of polarization of partially polarized, partially coherent pulsed electromagnetic beams propagating through anisotropic turbulence [J]. Optical engineering, 2019, 58(11): 116107-1–116107-11.
69. Moreno J, Castaeda R. Modeling of non-paraxial electromagnetic wave fields with partially spatially correlated amplitude and polarization [J]. Journal of Optics, 2021, 23(7): 075601–0756010.
70. Zhang Yixin, Song Zhengfang. Fluctuations in the propagation direction of laser beams in turbulent atmosphere [J]. Acta Optica Sinica, 1986(12):1111–1118.
71. Liu Weihui, Wu Jian. The effect of partially coherent beam on scintillation through strong turbulence [J]. Optics & Optoelectronic Technology, 2004,31(1):30–33
72. Feng Tao, Yao Min, Lin Qiang. Polarization properties of partially coherent light in random media [J]. Journal of Optoelectronics·Laser, 2004, 15(8): 990–993
73. Zhang Yixin, Tao Chunkan. The arrival-angle-fluctuation of Gaussian Schell beam propagating in turbulent atmosphere [J]. Acta Photon Sinica, 2005, 34(3):424–428.
74. Zhang Yixin. Comparison of two short-term propagation theories of beam in turbulent atmosphere [J]. Infrared and Laser Engineering, 2006, 35 (Suppl): 394–399.
75. Shi Lifen, Pu Jixiong, Chen Ziyang. Polarization changes of partially coherent electromagnetic beams propagating in turbulent media [J]. High Power Laser and Particle Beams, 2006, 18(8): 1271–1278

76. Gao Chong, Yu Siyuan, Ma Jing, et al. Beam wander in strong turbulence region [J]. High Power Laser and Particle Beams, 2006, 18(10):1597–1601.
77. Wang Chunyang, Li Jinshi. Detection of laser spot wander [J]. Applied Optics, 2007, 28(2): 205–209.
78. Qian Xianmei, Zhu Wenyue, Rao Ruizhong. Numerical simulation analysis of turbulence effect of atmospheric slant propagation of ground-to-air laser [J]. Infrared and Laser Engineering, 2008, 37(5):787–792.
79. X. Chu, Z. Liu, Y. Wu. Propagation of a general multi-Gaussian beam in turbulent atmosphere in a slant path [J]. J. Opt. Soc. Am. A., 2008, 25(1): 74–710.
80. Wang Tao, Pu Jixiong, Chen Ziyang, Shi Lifen. Polarization changes of partially coherent beams propagating in turbulent atmosphere [J]. Journal of Huaqiao University, 2008(02): 198–202.
81. Yang Ailin, Li Jinhong, Lv Baida. Comparative study of beam width expansion and angle spreading in atmospheric turbulence [J]. Acta Physica Sinica 2009, 58(04):2451–2460.
82. Yang Fan, Ji Xiaoling. Effect of atmospheric turbulence on the propagation of partially coherent electromagnetic flat-topped beam [J]. Acta Photon Sinica, 2009, 38(11):2948–2952.
83. Qian Xian-Mei, Zhu Wen-Yue, Wang An-Ting, Gu Chun, Rao Rui-Zhong. Numerical Simulation for Coherent and Partially Coherent Beam Propagation through Atmospheric Turbulence [J]. Chinese Physics Letters, 2010, 27(4): 044214–044219.
84. Wu GH, Luo B, Yu S ... et al. The propagation of electromagnetic Gaussian-Schell model beams through atmospheric turbulence in a slanted path [J]. Journal of optics A Pure and applied optics, 2011, 13(3):0357061–0357068.
85. Chen Kai, Cao Jianqiu, Lu Qisheng. Research progress on the coherence and polarization of light [J]. Optical Technique, 2011, 37(5):607–612
86. Zhang Xiaoxin, but right, Zhang Bin. Spreading of partially coherent beam propagating along a slant path in atmosphere [J]. Acta Optica Sinica, 2012, 32(12):8–16.
87. LI Ya-qing, WU Zhen-sen. Characteristics of a partially coherent Gaussian Schell-model beam propagating in slanted atmospheric turbulence [J]. Chinese Physics B, 2012, 21(05):237–244
88. Duan M L, Li J H, Wei J L. Influence of different propagation paths on the propagation of laser in atmospheric turbulence [J]. Opt. elec. Let. 2013, 9(6): 0477–0480.
89. Fei W, Cai Y, Eyyuboglu H T, et al. Average Intensity and Spreading of Partially Coherent Standard and Elegant Laguerre-Gaussian Beams in Turbulent Atmosphere [J]. Progress In Electromagnetics Research, 103 (2010):33–56.
90. Fang Guijuan, Lin Huichuan, Pu Jixiong. Generation and transmission of Besinc coherent vortex beam [J]. Chinese Journal of Lasers, 2015, 42(07):299–305
91. Hou Wenbo, Ma Dongdong. Research on transmission characteristics of partially coherent optical communication system [J]. Shipboard Electronic Countermeasure, 2015, 38(03): 47–50.
92. Zhang Yao, Chang Chengcheng, Chen Xudong, Pu Jixiong, Chen Ziyang, Lin Zhili. Atmospheric turbulence transmission characteristics of Gaussian inhomogeneous partially coherent beam [J]. Acta Photon Sinica, 2017, 46(09):121–126.
93. Wang Fei, Yu Jiayi, Liu Xianlong, et al. Progress in the study of partially coherent beam propagation through turbulent atmosphere [J]. Acta Physica Sinica, 2018, 67(18):9–22.
94. Zhou Zhenglan, Yuan Yangsheng, Shu Jie, Xu Xiang, Qu Jun. Wander of partially coherent crescent-shaped beam in non-Kolmogorov spectra [J]. Laser Technology, 2019, 43(04): 143–149.
95. Zhang Ying, Ma Chaoqun, Zhu Shijun, et al. Suppression of the effect of distorted Laguerre-Gaussian correlated beam on atmospheric turbulence propagation [J]. Chinese Optics, 2021, 14(3):704–716.
96. Yang Ning, Zhao Liang, Xu Ying, Xu Yonggen. Propagation characteristics of partially coherent radially polarized twisted beam in inhomogeneous atmospheric turbulence [J]. Laser and Infrared, 2022, 52(08): 1167–1176.

97. Ke Xizheng, Deng Lijun. Partially coherent optical transmission theory in Optical Wireless Communication [M]. Beijing: Science Press, 2016.
98. Ke Xizheng, Han Meimiao, Wang Mingjun. Spreading and wander of partially coherent beam propagating along a horizontal-path in atmospheric turbulence [J]. Acta Optica Sinica, 2014, 34(11):78–82.
99. Ke Xizheng, Zhang Yu. Scintillation of partially coherent beam in atmospheric turbulence [J]. Acta Optica Sinica, 2015, 35(01):56–62.
100. Ke Xizheng, Han Meimiao, Wang Mingjun. Spreading and wander of partially coherent beam through atmospheric turbulence in a slant the slant path in atmospheric turbulence [J]. Acta Photon Sinica, 2015, 44(03):149–153.
101. Ke Xizheng, Wang Wanting. Intensity and expansion of partially coherent beam in slant and horizontal atmospheric turbulence [J]. Journal of Applied Sciences, 2015, 33(02):142–154.
102. Ke Xizheng, Wang Wanting. Expansion and angular spread of partially coherent beam propagating in atmospheric turbulence [J]. Infrared and laser engineering, 2015, 44(09): 2726–2733.
103. Ke Xizheng, Wang Jiao. Comparison of uplink and downlink propagation polarization characteristics of partially coherent beam in atmospheric turbulence [J]. Acta Physica Sinica, 2015, 64(22):148–155.
104. Wang Jiao. Intensity distribution and polarization characteristics of partially coherent beam in atmospheric turbulence [D]. Xi'an University of Technology, 2016.
105. Wang Jiao, Ke Xizheng. Speckle characteristics of partially coherent beam propagating in atmospheric turbulence [J]. Infrared and Laser Engineering, 2017, 46(07):135–142.
106. Ke Xizheng, Xue Yao. The arrival-angle-fluctuation of partially coherent beam in atmospheric turbulence [J]. Acta Photon Sinica, 2016, 45(12):30–34.
107. Ke Xizheng, Xue Yao. The influence of atmospheric turbulence scale on the transmission characteristics of partially coherent beam [J]. Acta Photon Sinica, 2017, 46(01):35–42.
108. Ke Xizheng, Wang Chaozhen. Intensity distribution of partially coherent vortex beams propagating in atmospheric turbulence [J]. Laser & Optoelectronics Progress, 2016, 53(11): 83–91.
109. Ke Xizheng, Zhang Ya. Research on the transmission characteristics of partially coherent array beam in Optical Wireless Communication systems [J]. Laser & Optoelectronics Progress, 2016, 53(10):38–410.
110. Ke Xizheng, Zhang Huanjie. Study on the spreading characteristics of partially coherent Gaussian pulse in atmospheric turbulence [J]. Laser & Optoelectronics Progress, 2016, 53(08):110–116.
111. Ke Xizheng, Wang Song. Intensity evolution of partially coherent Airy beam in atmospheric turbulence [J]. Acta Photon Sinica, 2017, 46(07):7–18.
112. Ke Xizheng, Zhang Lin. Beam spreading and wander of partially coherent Airy beam in atmospheric turbulence [J]. Acta Photon Sinica, 2017, 46(01):43–410.
113. Ke Xizheng, Wang Chaozhen. Intensity distribution of partially coherent off-axis vortex beam in atmospheric turbulence [J]. Acta Optica Sinica, 2017, 37(01):44–50.
114. Wang Jiao, Ke Xizheng, Wang Mingjun. Influence of source parameters and atmospheric turbulence on the polarization properties of partially coherent electromagnetic vortex beams [J]. Applied Optics. 2019, 58(24): 6486–6494
115. Wang Jiao, Ke Xizheng, Wang Mingjun. Detecting the topological charge of partially coherent electromagnetic vortex beams through the orientation angle of polarization [J]. Journal of Quantitative Spectroscopy & Radiative Transfer, 2019, 228: 11–16.
116. Belen'kii, M. S. Effect of residual turbulent scintillation and a remote-sensing technique for simultaneous determination of turbulence and scattering parameters of the atmosphere [J]. J. Opt. Soc. Am. A, 1994, (11):1150–1159.
117. X. Ji, X. Chen and B. Lü. Spreading and directionality of partially coherent Hermite-Gaussian beams propagating through atmospheric turbulence [J]. J. Opt. Soc. Am. A, 2008, 25(1): 21–29.

118. Duan Meiling, Li Jinhong, Wei Jilin et al. Spreading of partially coherent Hermite Gaussian beam through atmospheric turbulence in slat path [J]. High Power Laser and Particle Beams, 2013, 25(9):2252–2256.
119. Li X, Ji X. Angular spread and directionality of the Hermite-Gaussian array beam propagating through atmospheric turbulence [J]. Appl Opt, 2009, 48(22):4338–4348.

Progress in Research on Channel Equalization in Optical Wireless Communication

17

Contents

Introduction 560
Foreign Research Progress 560
Domestic Research Progress 564
Key Channel Equalization Technologies 568
Linear Equalization 569
Nonlinear Equalization 572
Summary 575
References 576

Abstract

In optical wireless communication, undesirable channel is the main factor affecting system reliability, and placing an equalizer with opposite channel characteristics at the receiving end can offset the negative impact of the channel on the signal, so as to correct the inter-code crosstalk problem caused by multipath delay and other factors, and enable the signal to be recovered and reconstructed at the receiving end. On the basis of summarizing the research progress of equalization techniques both domestically and internationally, the chapter compares and summarizes various typical channel equalization techniques for reference, including zero-forcing equalization, minimum mean square error equalization, decision feedback equalization, Volterra equalization, etc. In practical applications, appropriate equalizers can be selected based on specific requirements and equalization algorithm characteristics to achieve channel equalization. Finally, the development prospects of channel equalization technology in optical wireless communication are prospected, providing new ideas and reference value for future researchers to explore in this field.

Keywords

Optical wireless communication · Channel equalization · Linear equilibrium · Nonlinear equilibrium

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_17

Introduction

Optical wireless communication is a new communication method that combines the advantages of fiber-optic and microwave communications and does not require spectrum licensing. Therefore, it has a wide range of applications [1] and has become a cutting-edge topic in the communication field, gradually attracting the attention of industry professionals.

Channel equalization is an anti-fading measure that suppresses the reduction in system-transmission performance caused by multipath delays and other issues. Channel equalization in optical wireless communication can be divided into multiple structures, based on different methods. Channel equalization technology is divided into two categories: linear and nonlinear, depending on whether the equalizer output is used for feedback control.

The former mainly eliminates the distortion and attenuation generated in the channel by filtering the signal but has high sensitivity to noise and signal interference. The latter mainly eliminates the distortion and attenuation generated in the channel by nonlinear signal processing. It can effectively improve the quality and reliability of the signal transmission and is little affected by noise and signal interference. However, the computational complexity is high and more computational resources are required.

Typical linear equalization techniques include minimum mean square error equalization, zero-forcing equalization, etc. Typical nonlinear equalization techniques include decision feedback equalization, maximum-likelihood sequence estimation, and maximum-likelihood symbol detection. Different equalization methods are suitable for different channel environments and signal characteristics and must be selected and applied according to actual needs.

The optical wireless communication environment is more complex and uncertain than that of wired transmission. Thus, it is usually necessary to calibrate or compensate the channel at the receiving end, because of the negative impact of imperfect channel conditions, to ensure efficient signal transmission and a low bit-error rate (BER). Imperfect channel conditions have become the main obstacle restricting the development and application of optical wireless communication [1–6]; therefore, studying channel equalization technology is an important topic in the field of optical wireless communication.

Foreign Research Progress

In 1994, Uysal et al. [7] proposed an optical wireless communication system based on adaptive equalizers that could adaptively adjust the equalizer parameters to offset distortion and interference in the channel, thereby improving the communication quality and reliability. This was considered a classic research study on channel equalization technology in the field of optical wireless communication and gradually attracted widespread attention from scholars. Subsequently, stochastic gradient

(SG) training algorithms were proposed for blind equalization algorithms, such as recursive least squares algorithms.

In 1994, Kechriotis et al. [8] found an adaptive recurrent neural network equalizer with a small size and high performance suitable for high-speed channel equalization. The results showed that when the channel interference was relatively small, the equalizer had performance comparable to traditional linear filter-based equalizers. When the channel's transfer function had a spectral zero or severe nonlinear distortion, the equalizer's performance was several orders of magnitude better.

In 1995, Cha et al. [9] proposed a random gradient training algorithm applicable to all the free parameters of the network. They found that complex radial basis function networks (CRBFNs) could be used as part of the equalization solution for complex nonlinear channels. The simulation results showed that the CRBFN using the gradient descent algorithm was highly effective for channel equalization.

In 1999, Patra et al. [10] proposed a computationally efficient single-layer functional link artificial neural network (FLANN), for adaptive channel equalization in nonlinear channels. Among three neural network equalizer structures, FLANN performed the best in terms of convergence speed, bit-error rate, and computational complexity for linear and nonlinear channel models over a wide range of signal-to-noise ratios.

In 2002, Deng et al. [11] proposed a complex radial basis function neural network for quadrature amplitude modulation (QAM) signal equalization. The network utilized the sequential learning algorithm of the complex minimum resource allocation network (CMRAN), and the simulation results showed its performance in nonlinear channel balancing problems. CMRAN had superior performance in terms of the symbol-error rate and network complexity. Compared with existing methods, CMRAN reduced the network complexity and therefore required less time for balancing.

In 2005, Pandey [12] proposed a new blind equalization scheme using a recursive least squares algorithm. The output layer weights were trained using an improved version of the constant modulus algorithm cost function. For weight training in the hidden-layer neuron space, the proposed scheme enabled the equalizer to converge more quickly.

In 2005, Patra et al. [13] introduced a Chebyshev-based equalizer structure for adaptive channel equalization in 4-QAM signal constellations. Simulation results showed that, compared with equalizers based on a multilayer perceptron and the least mean squares algorithm, its single-layer structure provided a faster convergence speed.

In 2007, Özden et al. [14] proposed a model-based Volterra-type nonlinear communication channel decision feedback equalization method. During the design, the nonlinear decision feedback equalization problem was visualized as a linear multichannel equalization problem. The elements of the required signal vector were estimated using the sigmoid activation function, which improved the error probability of the proposed equalizer, thus realizing a polynomial perceptron equalizer. The computational complexity calculations and performance results of the proposed decision feedback equalizer were presented.

In 2007, Al Awami et al. [15] used particle swarm optimization (PSO) to design adaptive equalizers, studied the effectiveness of PSO in channel equalization, discussed different modifications to the original PSO technology, and proposed new modification schemes. The performances of various PSO algorithms were compared with those of traditional least mean squares algorithms. The simulation results showed that, compared with adaptive equalizers based on the least mean squares and other algorithms, the proposed PSO algorithm converged faster under almost the same misalignments.

In 2011, Acolatse et al. [16] proposed a modified form of single-carrier frequency-domain equalization for diffuse optical wireless (DOW) communication. They proposed three different methods of using single-carrier frequency-domain equalization in DOW communication and demonstrated that they exhibited a lower peak-to-average power ratio (PAPR) and better bit-error rate performance in the case of LED nonlinearity.

In 2012, Diana D.C et al. [17] proposed a channel equalization method using inertial weights based on fitness-based PSO. A modified inertial weight PSO algorithm effectively eliminated the impact caused by multipath propagation, co-channel interference, and channel noise, and greatly reduced the time required to reduce errors; thus, a channel with dynamic characteristics was suitable for reliable transmission.

In 2014, Das et al. [18] proposed an equalizer based on an artificial neural network trained using a PSO algorithm. The algorithm was used to optimize the number of layers, input and hidden neurons, and transfer-function types. A large number of simulations showed that the proposed equalizer performed better in all noise conditions than other equalizers based on artificial neural networks and neuro fuzzy equalizers.

In 2015, Panda et al. [19] proposed a neural network training method. The training used directed search optimization (DSO) as a trainer for neural networks. It was then applied to nonlinear channel equalization problems, and a new nonlinear channel equalization strategy was introduced. Numerous simulation results indicated that the proposed channel equalization method outperformed other artificial neural network equalizers, in terms of the mean squared error and bit-error rate.

In 2016, Kaur et al. [20] proposed a radial basis function network for the adaptive channel equalization of a coherent optical orthogonal frequency-division multiplexing (OFDM) system and compared the simulation results with those of an artificial neural network based on a multilayer perceptron, which also had good performance.

In 2017, Nanda et al. [21] produced an equalizer based on wavelet neural networks and weighted it using metaheuristics (a symbiotic-biology search algorithm). The performance of the proposed equalizer was compared with that of a wavelet neural network trained by cat swarm optimization, a clonal selection algorithm, particle swarm optimization, and a least mean squares algorithm. This demonstrated the robustness and superior performance of the equalizer in handling sudden error conditions.

In 2018, Dai et al. [22] developed a channel equalization method based on a set membership filter and a binormalized modified-norm algorithm. Its performance advantages in convergence speed and eliminating intersymbol interference were verified by comparing it with existing channel equalization methods. This algorithm was theoretically significant for channel equalization in high-speed free-space optical communication systems.

In 2019, Saidani et al. [23] proposed a new method for optimizing the neural network weights for channel equalization. The update strategy in the local search of the standard leapfrog algorithm was replaced with the update strategy of the directed search optimization algorithm. The proposed method was compared with standard PSO, an improved modified shuffled frog-leaping (MSF) algorithm, and a directed search-optimization algorithm to evaluate its performance in nonlinear channel equalization. The experiment showed that this method achieved the best results and outperformed other algorithms.

In 2019, Wu et al. [24] proposed a kernel recursive maximum Versoria-like criterion (KRMVLC) algorithm that considered the advantages of the logarithmic second-order error and symmetric maximum Versoria criterion (MVC) in a reproducing kernel Hilbert space (RKHS). The proposed KRMVLC algorithm was derived for nonlinear channel equalization (NCE) problems under different non-Gaussian interference. The results verified that KRMVLC was robust to non-Gaussian interference.

In 2020, Caciuaru et al. [25] proposed a new unsupervised neural network blind channel equalization algorithm that simultaneously considered linear and nonlinear noise intersymbol interference channels. The results were then extended to the joint equalization and decoding of low-density parity codes using iterative algorithms. Compared with the baseline algorithm, it demonstrated a significantly improved bit-error rate performance.

In 2020, Linsalata et al. [26] proposed the design of linear and nonlinear receivers for generalized frequency-division multiplexing (GFDM) transmissions using gradient diffusion film technology on frequency-selective channels. They analyzed and compared the maximum-likelihood, zero-forcing, and minimum mean squared error equalization schemes. A Monte Carlo simulation was used to measure the bit-error rate performance of linear and nonlinear receivers considering frequency- and non-frequency-selective channels. The numerical results indicated that this design improved performance compared to other methods.

In 2021, Uma et al. [27] introduced a new theory to distinguish the step sizes based on evolutionary programming algorithms. This method produced a high overall efficiency and did not require presetting the considerations involved in signal-based numerical characteristic solutions. The weighted-variance weighting-coefficient deviation depended on changes in the algorithm. The simulation results showed that this algorithm had good performance.

In 2022, Mohapatra [28] proposed a hybrid algorithm integrating the PSO algorithm and the gray wolf optimization algorithm to improve the flexibility of all algorithm variants and made use of appropriate weights, network topologies, and neural network transfer functions in channel equalization. Experimental simulations

showed that the performance of this equalizer was superior to that of existing neural network-based equalizers and neural fuzzy equalizers.

In 2022, Mohapatra et al. [29] proposed a new nonlinear channel-modeling method that used a radial basis function neural network to train a new cuckoo search algorithm for equalization in nonlinear channels. The experimental simulation results showed that the algorithm improved the equalization efficiency during the search process, and the equalizer performance of this nonlinear channel was superior to that of existing neural network-based equalizers.

In 2022, Badal et al. [30] developed a naturally inspired butterfly optimization algorithm and applied its training strategy to channel equalization. It was found that this equalizer performed better than previously known equalizers in terms of the bit-error rate and mean square error.

In 2022, Mohapatra et al. [31] proposed a new training strategy for channel equalization using a fuzzy firefly algorithm (FFA). By using an appropriate network topology and parameters, the training system provided stronger development and exploration skills, as well as the ability to solve local minimum problems. The simulation results showed that the proposed method validated a wide range of signal-to-noise ratios and outperformed existing neural network-based equalizers.

Research on optical wireless communication channel equalization has been ongoing since 1994, with a shift in focus from the initial linear channel equalization methods to nonlinear channel equalization methods to improve the communication quality. The research content also developed from theoretical derivations and simulations to practical applications. As research has gradually developed from traditional channel equalization methods to combining channel equalization in the field of neural networks, the performance of bit-error rates and other aspects has improved.

Domestic Research Progress

In 2003, Cai et al. [32] proposed an improved minimum mean square algorithm with good equalization performance for time-varying channels.

In 2005, Sun [33] proposed a wavelet neural network with a new initialization method and applied it to channel equalization. It was found that this method had better anti-nonlinearity ability than linear equalizers.

In 2006, Zhang et al. [34] proposed a channel equalization method using PSO to adaptively adjust the equalizer coefficients; this method also achieved good equalization results.

In 2007, Long [35] proposed a method based on linear precoding technology. Simulation experiments showed that this equalization method could effectively improve the BER performance of a system under a high signal-to-noise ratio.

In 2007, Mi et al. [36], in order to further verify the ability of decision feedback equalizer to eliminate intersymbol interference, simulated in Matlab and compared it with the learning curve of linear equalizer, and found that the equalization effect of decision feedback equalizer is better than that of linear equalizer.

In 2008, Chen et al. [37] proposed a decision feedback equalizer based on unscented Kalman filtering (UKF), which reduced the mean squared error of the system.

In 2008, Xu et al. [38] proposed a new constant modulus algorithm (CMA) with a variable step size that could effectively reduce system errors.

In 2010, Li [39] proposed a variable step-size algorithm and a variable step-size normalization algorithm based on decorrelation. Experimental simulations showed that this algorithm effectively solved the slow convergence speed problem of the variable step-size algorithm, and the algorithm performance was also significantly improved.

In 2010, Zhao et al. [40] proposed a second-order Volterra adaptive algorithm based on lattice filtering. The simulation results showed that the proposed algorithm had a faster convergence speed and better steady-state performance for both linear and nonlinear channels than the domain second-order Volterra-series adaptive equalizer under nonlinear channels.

In 2011, Wang et al. [41] proposed a channel equalization method based on turbo iterations. The simulation results showed that this method could significantly reduce the error rate of communication systems.

In 2011, Wang [42] realized an unconstrained linear equalization algorithm using a factor graph. The performance of this algorithm was superior to that of traditional equalization algorithms, and the computational difficulty was reduced to a linear function of the data length.

In 2012, Wang [43] proposed a modified constant modulus algorithm (MCMA) based on the constellation matching error, the decision-guided dual-mode MCMA algorithm, and the stop-and-go-based MCMA algorithm. These blind equalization algorithms reduced the intersymbol interference and improved the convergence speed.

In 2013, Zhang [44] proposed a single-carrier minimum mean square error equalization algorithm based on iterative thinking, after considering two types of interference: intersymbol interference (ISI) and noise interference. This algorithm considered noise interference based on a decision feedback algorithm and was compared with traditional equalization algorithms. The performance curves of the minimum mean square error residual ISI cancellation (MMSE RISIC) decision feedback equalization algorithm were simulated and compared with the traditional algorithms. It was found that the combined algorithm outperformed the latter two in terms of signal-to-noise ratio and bit-error rate.

In 2014, Yang et al. [45] proposed an improved equalization method for sparse multipath channels and a greedy algorithm to obtain an approximate optimal solution. The simulation results showed that this method significantly improved the system error rate performance compared to the traditional minimum mean square linear method and achieved a power margin of approximately 10 dB at a signal-to-noise ratio of 15 dB.

In 2015, Luo [46] proposed an MMSE equalization algorithm that combined symbol reconstruction at the receiver and compared it with zero-forcing equalization, minimum mean square error equalization, decision feedback equalization, and

maximum-likelihood sequence estimation equalization analysis. It had a lower bit-error rate and an improved signal reuse rate.

In 2016, Li [47] verified the excellent performance of an adaptive algorithm based on the reproducing kernel Hilbert space (RKHS) and found that the algorithm was suitable for the time-varying nonlinear channel equalization field.

In 2016, Wang [48] studied a frequency-domain equalization method based on channel and noise estimation, which compensated for the channel in the frequency domain and compared it with frequency-domain zero-forcing equalization and frequency-domain minimum mean square error equalization. Simulation results showed that the algorithm could effectively improve the performance of atmospheric laser communication systems in the frequency domain.

In 2017, Chen et al. [49] proposed a neural network-based channel equalization algorithm. Compared with the recursive least square (RLS) and multilayer-perceptron algorithms, the symbol-error rates of the algorithm for quadrature phase-shift keying (QPSK) recovery were reduced by 20% and 5%, respectively.

In 2017, Sun et al. [50] constructed a nonlinear equalizer based on a kernel method under stable distribution noise. Simulation results showed that this algorithm reduced the steady-state error while the convergence speed remained unchanged. The nonlinear channel distortion under stable distribution noise had a better effect.

In 2018, Li [51] proposed an equalization algorithm based on adaptive bit modulation that could effectively improve the signal output gain and reduce the communication error rate.

In 2019, Guo [52] proposed a new chromaticity dispersion-compensation algorithm, three new polarization-mode dispersion-equalization algorithms, and a cascaded multimode algorithm with adaptive momentum terms. All three algorithms had better bit-error rate performance and equalization effects than the traditional algorithms.

In 2019, Ma et al. [53] proposed an equalization algorithm based on compressed sensing. This algorithm had a similar robustness and better equalization performance than traditional linear equalization algorithms.

In 2020, Yang et al. [54] used the minimum mean squared error criterion to design a linear equalizer and tested it on a universal software radio peripheral (USRP) platform. The experimental results showed that the error rate remained around 10^{-5} after equalization, and the stability and reliability were improved to some extent.

In 2020, Li et al. [55] built an adaptive equilibrium system model and analyzed the improved normalized minimum mean square algorithm with an orthogonal correction factor. They simulated and compared the traditional algorithm and the improved algorithm. The proposed algorithm had better convergence characteristics.

In 2020, Lü [56] used the least mean square (LMS) algorithm in a sliding autoregressive model to verify the convergence performance of the LMS algorithm through simulation under second-order conditions. It achieved better equalization performance in the finite impulse response (FIR) structure under QPSK modulation channels.

In 2021, Baiguo et al. [57] jointly optimized channel estimation, noise power estimation, and channel equalization to propose a single-carrier frequency-domain

equalization algorithm. The simulation results showed that this algorithm had a lower bit-error rate compared to traditional algorithms.

In 2021, Chen et al. [58] proposed a crosswalk detection network (CD net) channel equalization method based on artificial neural networks. The simulation results showed that, compared with traditional equalization methods, this algorithm provided a better bit-error rate performance.

In 2022, Wang [59] proposed an equilibrium algorithm based on complex Gaussian-process regression, which applied the idea of kernel recursion to complex Gaussian processes. This algorithm exhibited better equilibrium performance than other traditional equilibrium algorithms.

In 2022, Marina et al. [60] proposed an improved *k*-means algorithm based on the nonlinear transmission equation of the Volterra series using the traditional *k*-means algorithm to achieve channel equalization in orthogonal frequency-division multiplexing/offset quadrature amplitude modulation systems (OFDM/OQAM). Experimental comparisons and simulations were conducted before and after the improvement, indicating that the improved *k*-means algorithm required significantly fewer iterations than the traditional algorithms, had a simple structure, located the centroid faster, and significantly reduced linear and nonlinear distortions.

In 2022, Xie [61] used a turbo equalization algorithm to process the receiver signal and added a loop-head structure to the traditional minimum equalization error algorithm, which improved the performance and efficiency of the original algorithm in the frequency domain. The experimental results showed that the error rate performance of the improved algorithm improved by approximately 5 dB, compared to that of the previous algorithm.

In 2022, Chen [62] designed a channel equalization method based on a long short-term memory (LSTM) neural network in combination with an LSTM network. The proposed scheme was experimentally verified and analyzed. It was found that under medium and strong turbulence, the performance of the LSTM equalizer maintained a stable improvement in the bit-error performance of the system and remained stable in the face of a turbulent random environment.

In 2022, Menda [63] proposed a method based on a backpropagation (BP) neural network algorithm. Simulation experiments were conducted using three types of equalizers: the recursive least squares algorithm, minimum mean square algorithm, and BP neural network algorithm. The simulation results showed that the algorithm constructed using the BP neural network had the best equalization effect, could adapt to real-time channel changes in visible light systems, and compensated for channel characteristics. The error rate reached 10^{-6}.

In 2022, Shi [64] improved structure based on neural networks and compared the rectified linear unit (ReLU), tanh, leaky ReLU, and other activation functions. They found through simulation experiments that the neural network adaptive equalization scheme constructed by the maxout activation function had a better effect and the ability to effectively nonlinearly compensate for the received signal, resulting in a faster convergence speed and smaller mean square error.

In 2022, Chen et al. [65] proposed a zero-value attraction sparse-control proportional minimum symbol-error rate decision feedback equalization (ZA-SC-PMSER

DFE) algorithm to improve the convergence of complex underwater acoustic channel equalization algorithms and reduce their bit-error rates. The convergence speed and other performance metrics were superior to other algorithms, such as ZA-SC-PMSER DFE without a phase-locked loop and ZA-SC-PMSER DFE with a phase-locked loop, based on the minimum mean square error criterion.

In 2022, Duan et al. [66] showed that three deep learning channel equalization algorithms had lower bit-error rates and better overall performance than traditional channel equalization algorithms. Among them, the algorithm based on convolutional neural networks exhibited the best performance. Under this algorithm, when the input-sequence symbol length was 2000, the bit-error rate reached the minimum value, which was equivalent to 3.65 times the channel length.

Various channel equalization algorithms in optical wireless communication are interrelated, and each algorithm is proposed and improved, based on the shortcomings of existing algorithms. However, the new algorithm also has limitations. Thus far, nearly 20 channel equalization methods have been presented for optical wireless communications, and they still have unresolved problems. In future research on channel equalization, additional algorithms will be proposed to suppress the performance losses caused by imperfect channels during signal transmission.

Key Channel Equalization Technologies

The main purpose of channel equalization technology is to eliminate intersymbol crosstalk and compensate for channel distortion, thereby improving the transmission performance of communication systems.

Two main channel equalization techniques are used in optical wireless communication: linear and nonlinear, with various structures. The structures are illustrated in Fig. 17.1. The equalizer is essentially a reciprocal filter for the channel transfer function. According to the different channel characteristics and application scenarios, suitable linear or nonlinear channel equalization methods can be selected to

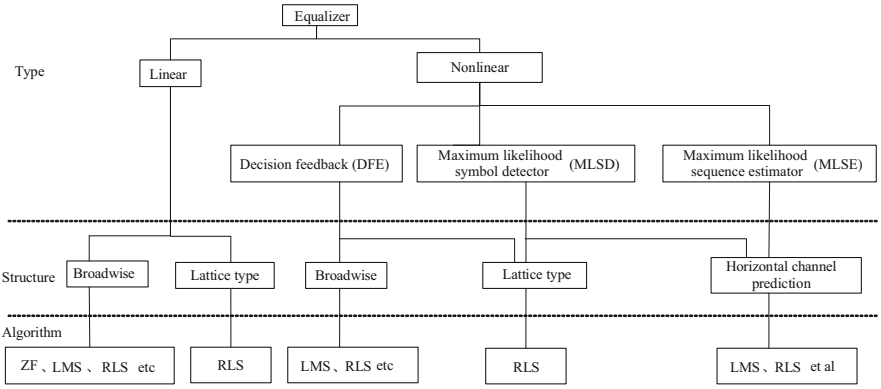


Fig. 17.1 Common equalizer structures [47]

improve the performance and reliability of optical wireless communication. This chapter introduces several typical linear and nonlinear equalization methods.

Linear Equalization

In optical wireless communications, linear channel equalization can eliminate signal distortion and interference through a linear transformation, thereby equalizing the channel. Typical linear equalization algorithms include zero-forcing and minimum mean square error equalization.

The basic structure of a linear equalizer is represented by a transverse filter, as shown in Fig. 17.2. The transverse filter mainly consists of $2N$ delay units T_s, $2N + 1$ tap coefficients, a tap-gain adjustment algorithm, and an adder. Each tap coefficient can be adjusted, and different tap coefficients can be set according to the channel characteristics to compensate for the loss caused by intersymbol interference in the channel.

Assuming that the sequence of input signal sampling values is $\{x_k\}$, after passing through $2N$ delay units T_s, each sampling value is multiplied by its corresponding tap coefficient C_N to obtain the balanced output signal sampling value sequence $\{y_k\}$. The balanced output is used as the input of the tap-gain adjustment algorithm to adjust the numbers of the tap coefficients, such that equalization can be achieved at different sampling values. The output signal sampling value sequence $\{y_k\}$ is calculated using the following formula:

$$y_k = \sum_{i=-N}^{N} C_i x_{k-i}, \quad k = -2N, \cdots, 0, \cdots, +2N, \tag{17.1}$$

where C_{-N}, C_{-N+1}, $\cdots$, C_N are the tap coefficients of the equalizer, x_k is the input of the equalizer, and y_k is the output of the equalizer.

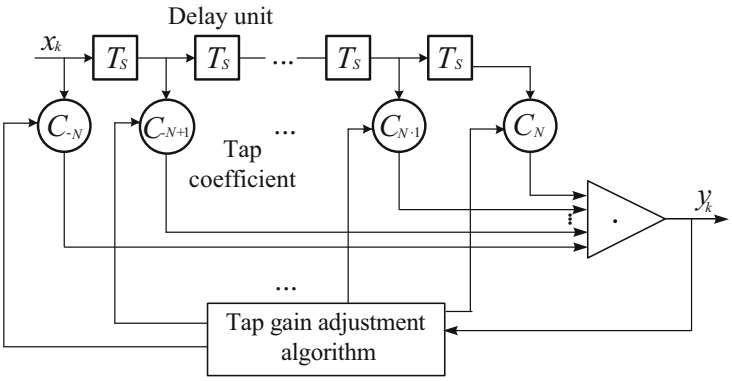


Fig. 17.2 Transverse filter structure [1]

Designing a transverse filter is a key approach to linear equalization. Two commonly used methods are the zero-forcing equalization algorithm, based on the minimum peak distortion, and the mean square error equalization algorithm based on the minimum mean square error [1]. The key to equalization is to design appropriate equalization algorithms to adaptively adjust the tap coefficients based on the channel characteristics.

Zero-Forcing Equalization Algorithm

Lucky proposed a linear equalization filter for zero-forcing equalization in 1965 [1]. The minimum peak distortion is an important criterion for measuring the equalization effect, and its calculation formula is

$$D = \frac{1}{y_0} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} |y_k|. \tag{17.2}$$

When $k \neq 0$, D represents the ratio of the sum of y_k absolute values at all sampling times to the sampling time y_0 when $k = 0$. According to Eq. (17.2), the input peak error is defined as

$$D_0 = \frac{1}{x_0} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} |x_k|. \tag{17.3}$$

The output peak error is defined as

$$D = \frac{1}{y_0} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} |y_k|. \tag{17.4}$$

When the input peak error $D_0 < 1$, the minimum value of the output peak error function appears near $\begin{cases} y_0 = 1 \\ y_k = 0, 1 \leq |k| \leq N \end{cases}$; the $2N + 1$ linear equation that the tap coefficient must meet is known from Eq. (17.5):

$$\begin{cases} \sum_{i=-N}^{N} C_i x_{k-i} = 0, k = \pm 1, \pm 2, \cdots, +N \\ \sum_{i=-N}^{N} C_i x_{-i}, k = 0 \end{cases}. \tag{17.5}$$

Here, it is written in matrix form:

$$\begin{bmatrix} x_0 & x_{-1} & \cdots & x_{-2N} \\ \vdots & \vdots & & \vdots \\ x_N & x_{N-1} & \cdots & x_{-N} \\ \vdots & \vdots & & \vdots \\ x_{2N} & x_{2N-1} & \cdots & x_0 \end{bmatrix} \begin{bmatrix} C_{-N} \\ C_{-N+1} \\ \vdots \\ C_0 \\ \vdots \\ C_{N-1} \\ C_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{17.6}$$

When the input sequence $\{x_k\}$ is known, the tap coefficient C_N of the equalizer can be calculated according to the matrix equation system of Eq. (17.6). At a later stage, the tap coefficient can also be adjusted to force the output sampling values $\{y_k\}$ of the equalizer to zero, optimizing the equalization effect [1].

The zero-forcing algorithm is suitable for channels with light multipath fading, mainly because it ignores the additive noise in the channel. Therefore, the tap coefficient value obtained when the channel contains additive noise is not optimal; however, its implementation is relatively simple.

Minimum Mean Square Error Equalization Algorithm

The core concept of this method is to equalize the channel by minimizing the mean square error (MSE) between the received and original signals. The channel equalization effect can be measured using the mean square distortion; its calculation formula is

$$e^2 = \frac{1}{y_0} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} {y_k}^2, \tag{17.7}$$

where y_k represents the sample value after equilibrium is reached. The transmission sequence is set to $\{a_k\}$, and after equalization, it is output to y_k, where the error signal is

$$e_k = y_k - a_k. \tag{17.8}$$

At this point, the mean square error is defined as

$$\overline{e^2} = E(y_k - a_k)^2. \tag{17.9}$$

Substituting Eq. (17.7) into Eq. (17.9),

$$\overline{e^2} = E\left(\sum_{-N}^{N} C_i x_{k-i} - a_k\right)^2. \tag{17.10}$$

The mean square error $\overline{e^2}$ is a function with each tap coefficient as the independent variable. It can be adjusted to achieve equalization by adjusting each tap coefficient of the equalizer such that they satisfy

$$\frac{\partial \overline{e^2}}{\partial C_i} = 0, \quad i = \pm 1, \pm 2, \cdots, \pm N. \tag{17.11}$$

According to Eqs. (17.10) and (17.11),

$$\frac{\partial \overline{e^2}}{\partial C_i} = 2E[e_k x_{k-i}] = 0, \quad i = \pm 1, \pm 2, \cdots, \pm N. \tag{17.12}$$

In the equation,

$$e_k = y_k - a_k = \sum_{i=-N}^{N} C_i x_{k-i} - a_k. \tag{17.13}$$

According to Eq. (17.12), when the cross-correlation function between the error signal and the input sampling value is zero, the derivative of the mean square error is zero. At this point, the error reaches its minimum value, and the resulting tap coefficient is the best. The least mean squares algorithm can not only equalize the intersymbol interference caused by multipath transmission delay but also the influence of channel additive noise on the signal; however, this method cannot completely eliminate the intersymbol interference [1].

Nonlinear Equalization

In nonlinear channels, the interference and distortion of the signals transmitted through the channel are nonlinear. Linear channel equalization methods cannot eliminate the nonlinear distortion and interference, and often do not achieve satisfactory equalization results. Therefore, nonlinear channel equalization methods should be used to equalize the channel. Nonlinear channel equalization can eliminate signal distortion and interference through nonlinear transformations, thereby equalizing the channel. Typical nonlinear equalization algorithms include decision feedback equalizers and neural network channel equalizers.

Decision Feedback Equalizer

The decision feedback equalization structure was formally proposed by Austin in 1969 [1]. Its implementation complexity is low, which reduces the impact of channel intersymbol interference and additive Gaussian white noise on the system, and improves the system performance to a certain extent.

The basic principle of decision feedback equalization is to subtract the signal values received at the current sampling time from the signal values received at the signal-receiving end at other sampling times to eliminate the intersymbol interference

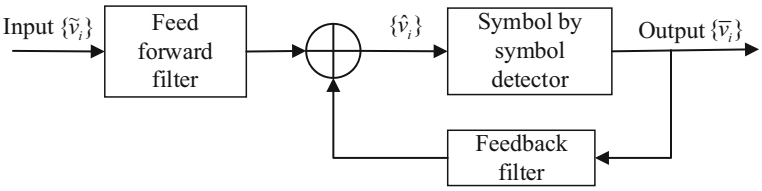


Fig. 17.3 Decision feedback equalization structure [31]

generated by other sampling times on the current sampling time and equalize and recover the data.

Decision feedback equalization is achieved by cascading a feedforward filter and a feedback filter, as shown in Fig. 17.3. Assuming that the input signal sequence is $\{\tilde{v}_i\}$, the input of the feedforward filter is $\{\tilde{v}_i\}$; and the input of the feedback filter is the output $\{\overline{v}_i\}$ of the decision feedback device. The output of the DFE is

$$\hat{v}_i = \sum_{k=-K_1}^{0} c_k \tilde{v}_{i-k} + \sum_{k=1}^{K_2} c_k \overline{v}_{i-k} \qquad i = \ldots, -2, -1, 0, 1, 2, \ldots. \tag{17.14}$$

In the equation, $\{c_k\}$ is the equalizer tap coefficient, with a total of $K_1 + K_2 + 1$ tap coefficients, $K_1 + 1$ for the feedforward part and K_2 for the feedback part. After adding the output of the feedforward filter and the output of the feedback filter, symbol-wise decisions are made, and the symbol decision result is used as the output and input of the feedback filter.

If the tap coefficient is ideal, the combination of the two can cancel the intersymbol interference between symbols in the channel, eliminate the interference of previous symbols on the current symbol vector, and theoretically eliminate all intersymbol interference, thus equalizing the channel [34].

Maximum-Likelihood Criterion

The maximum-likelihood symbol detector and maximum-likelihood sequence estimator are both nonlinear equalizers that can, to some extent, suppress the noise interference generated in the channel at the receiving end of the communication system. Their principle is based on the maximum-likelihood criterion, which selects the transmission signal that is most likely to generate the received signal under the given conditions.

Maximum-likelihood estimation (MLE) was proposed by British statistician Ronald Fisher in 1922 [67]. The basic idea is to estimate the parameter using known observations without any prior knowledge of the parameter so that the selected sample is most likely to appear in the selected population.

Let a represent N observations of $x_1, \ldots, x_k$ random variables, and $\{f(x_1, \cdots, x_N| \theta), \theta \in \theta\}$ represent a given parameter θ. Assume that the joint conditional probability density function exists and is bounded, and consider

unknown parameters θ. When considering the joint conditional distribution density function c as a true parameter, θ, it is often called a likelihood function. Maximum-likelihood estimation is recorded as follows:

$$\widehat{\theta}_{\mathrm{ML}} = \arg\max_{\theta\in\theta} f(x_1, \cdots, x_N|\theta). \tag{17.15}$$

The equilibrium structure of the maximum-likelihood symbol detector is usually a decision-maker used to map the received signal to the most likely sending symbol. The formula can be expressed as

$$\widehat{s} = \arg\max_{s_i} P(r|s_i), \tag{17.16}$$

where $\widehat{s}$ represents the estimated transmission symbol; r represents the received signal; s_i represents the actual sent symbol; and $P(r \mid s_i)$ represents the probability of receiving r when sending symbol s_i.

The equilibrium structure of a maximum-likelihood sequence estimator is typically a filter used to estimate the transmitted symbol sequence. The formula can be expressed as

$$\widehat{\mathbf{s}} = \arg\max_{\mathbf{s}} P(\mathbf{r}|\mathbf{s}), \tag{17.17}$$

where $\widehat{\mathbf{s}}$ represents the estimated sequence of transmission symbols; $\mathbf{r}$ represents the received signal sequence; $\mathbf{s}$ represents the actual sequence of sent symbols; and $P(\mathbf{r} \mid \mathbf{s})$ represents the probability of receiving $\mathbf{r}$ when sending symbol sequence $\mathbf{s}$.

Both the maximum-likelihood symbol detector and maximum-likelihood sequence estimator are nonlinear equalizers designed using statistical principles, which can further improve the performance of optical wireless communication systems.

Channel Equalization Algorithms Based on Neural Networks

Channels are modeled and predicted using neural network models to equalize the channels. Neural networks exhibit strong nonlinear characteristics and play important roles in signal analysis and processing. Compared with the equalization effect of traditional linear equalizers, a system that uses neural networks to equalize channels has a lower bit-error rate.

For example, the channel equalization of a support vector machine (SVM) utilizes the SVM network model to model and equalize the channel. Particle swarm optimization (PSO) channel equalization optimizes the channel parameters through PSO. Genetic algorithm (GA) channel equalization optimizes the channel parameters through genetic algorithms. These nonlinear channel equalization methods are widely used in optical wireless communication.

Compared with linear channel equalization methods, nonlinear channel equalization methods can better adapt to more complex channel characteristics and improve the performance of optical wireless communication. However, their computational complexity is relatively high, requiring additional computational resources and time.

Appropriate methods must be selected according to specific application scenarios and channel characteristics.

Summary

This chapter provided a brief overview of the foreign and domestic research status of channel equalization technology in optical wireless communication. It provided a brief introduction to typical channel equalization technologies in two categories—nonlinear and linear—and analyzed the advantages and disadvantages of linear and nonlinear equalization methods, as shown in Table 17.1.

It was found that, in the case of low channel noise and a narrow signal spectrum, the performance of linear channel equalization was better than that of nonlinear channel equalization, whereas in the case of high channel noise and a wide signal spectrum, the performance of nonlinear channel equalization was better than that of linear channel equalization.

The prospects of channel equalization are very broad, and future research will likely include the following:

1. Research should dig deeper into blind equalization algorithms. Currently, many algorithms require knowledge of the training sequence to determine the tap coefficients, resulting in low equalization efficiency. To avoid this issue, blind equalization algorithms use the transmitted signal characteristics to determine the equalizer tap coefficients.
2. Research should focus on the cross integration of multiple technologies, design new generalized activation functions to improve the ability of neural networks, realize neuro fuzzy systems, develop algorithms and neural network structures, and balance complex time-varying channels with faster convergence and simpler, more efficient architectures.
3. Researchers should study new metaheuristic algorithms, note the importance of swarm intelligence, find inspiration from natural communities, flocks, herds, and

Table 17.1 Advantages and disadvantages of linear and nonlinear equalization methods

	Advantages	Disadvantages
Linear equalization	1. Simple implementation and fast calculation speed 2. Better effect on low channel noise 3. Suitable for narrow spectrum widths	1. Unsuitable for high channel noise 2. Channel characteristics must be known, which is difficult with complex channels 3. Cannot handle nonlinear distortion issues
Nonlinear equalization	1. Can handle nonlinear distortion issues, with good results for large nonlinear channel distortion 2. Can adaptively handle channel characteristics and equalize complex channels	1. Complex implementation and slow calculation speed 2. Poor effect on high channel noise 3. Only suitable for situations with wide spectra

schools, and study new metaheuristic algorithms based on a single problem, such as a slow convergence rate or based on groups.

4. Research should focus on addressing the shortcomings of short transmission distances and narrow coverage, identify transmission wavelengths with stronger penetration, and enable them to overcome obstacles and achieve longer transmission distances. Technology should be selected and optimized, based on specific application scenarios.
5. In terms of hardware, more attention should be paid to reducing the complexity of hardware devices, amount of system computation, and power consumption, and improving the convergence rate of signal errors, thus compensating for system performance.

References

1. Ke Xizheng, Deng Lijun Optical Wireless Communication [M] Beijing: Science Press, 2016: 117–120
2. M. Uysal, H. Haas, Adaptive Equalization for Optical Wireless Communications [J]. in IEEE Journal on Selected Areas in Communications, 1994,6(3):134–148.
3. G. Kechriotis, E. Zervas and E. S. Manolakos, Using recurrent neural networks for adaptive communication channel equalization [J]. IEEE Transactions on Neural Networks, 1994,3(5): 267–279.
4. Inhyok Cha and Saleem A. Kassam. Channel Equalization Using Adaptive Complex Radial Basis Function Networks [J]. IEEE Journal on Selected Areas in Communications, 1995, 13(1): 122–131.
5. Patra J C, Pal R N, Baliarsingh R, Panda G. Nonlinear channel equalization for QAM signal constellation using artificial neural networks [J]. IEEE transactions on systems,1999,29(2): 225–236.
6. Deng Jianping,Sundararajan Narasimhan,Saratchandran P. Communication channel equalization using complex-valued minimal radial basis function neural networks [J]. IEEE transactions on neural networks,2002,13(3):122–132.
7. Pandey R. Fast blind equalization using complex-valued MLP [J]. Neural processing letters, 2005, 21(2): 215–225.
8. Patra J C, Wei B P, Chaudhari N S, et al. Nonlinear channel equalization with QAM signal using Chebyshev artificial neural network [J]. IEEE International Joint Conference on Neural Networks. IEEE, 2005,31(2):145–153.
9. M. Tahir Özden, Ahmet H. Kayran. Adaptive multichannel decision feedback equalization for volterra type nonlinear communication channels [J]. AEUE - International Journal of Electronics and Communications,2007,62(6):322–338.
10. T. Al-Awami, W. Saif, A. Zerguine, A. Zidouri and L. Cheded, An adaptive equalizer based on particle swarm optimization techniques [J]. 2007 9th International Symposium on Signal Processing and Its Applications, Sharjah, United Arab Emirates, 2007,(15)23: 1–4.
11. Kodzovi Acolatse, Yeheskel Bar-Ness. Novel Techniques of Single-Carrier Frequency-Domain Equalization for Optical Wireless Communications [J]. EURASIP Journal on Advances in Signal Processing,2011,(12)23:10-1.
12. Diana D.C, Rani S.P, Nithya T.R. Adaptive Inertia Weight Particle Swarm Optimization For Linear and Nonlinear Channel Equalization [J]. International Journal of Scientific Research,2012,3(4):346–358.
13. Gyanesh Das and Prasant Kumar Pattnaik and Sasmita Kumari Padhy. Artificial Neural Network trained by Particle Swarm Optimization for non-linear channel equalization [J]. Expert Systems With Applications, 2014, 41(7): 3491–3496.

14. Sunita Panda and Pradyumna K. Mohapatra and Siba Prasada Panigrahi. A new training scheme for neural networks and application in non-linear channel equalization [J]. Applied Soft Computing Journal, 2015, 27(9) :47–52.
15. Gurpreet Kaur and Gurmeet Kaur. A Radial Basis Function Network for Adaptive Channel Equalization in Coherent Optical OFDM Systems [J]. Research Cell: An International Journal of Engineering Sciences, 2016, 20(0): 121–125.
16. Satyasai Jagannath Nanda and Nidhi Jonwal. Robust nonlinear channel equalization using WNN trained by symbiotic organism search algorithm [J]. Applied Soft Computing, 2017, 57(0): 197–2010.
17. Dai Z, Hao S, Zhao Q, et al. Differential phase-shift keying and channel equalization in free space optical communication system [J]. Optical Engineering, 2018, 57(1):015107–015108.
18. Samir Saidani and A. Moussaoui and M. Ghadjati. Channel Equalization Based on SFLA and DSO Trained Artificial Neural Network [J]. Telecommunications and Radio Engineering, 2019, 78(17):189–194.
19. Qishuai Wu and Yingsong Li and Wei Xue. A Kernel Recursive Maximum Versoria-Like Criterion Algorithm for Nonlinear Channel Equalization [J]. Symmetry, 2019, 11(9): 1067–1068.
20. Avi Caciularu and David Burshtein. Unsupervised Linear and Nonlinear Channel Equalization and Decoding using Variational Autoencoders [J]. IEEE Transactions on Cognitive Communications and Networking, 2020, 13(99): 1-1.
21. Francesco Linsalata and Atul Kumar and Maurizio Magarini. Design of DGT-based linear and non-linear equalizers for GFDM transmission [J]. Physical Communication, 2020, 42(12): 255–256.
22. Uma K. et al. A New Variable Step Size Least Mean Square Evolutionary Algorithm for Channel Equalization [J]. Journal of Computational and Theoretical Nanoscience, 2021, 18(3):935–941.
23. Kumar Mohapatra Pradyumna et al. A Novel Application of HPSOGWO Trained ANN in Nonlinear Channel Equalization [J]. Springer Nature Singapore, 2022:159–174.
24. Mohapatra Pradyumna Kumar et al. A Novel Cuckoo Search Optimized RBF Trained ANN in a Nonlinear Channel Equalization [J]. Springer Nature Singapore, 2022,13(2): 189–203.
25. Acharya Badal et al. A novel approach for BOA trained ANN for channel equalization problems [J]. Journal of Information and Optimization Sciences, 2022, 43(8): 2121–2130.
26. P. K. Mohapatra, S. K. Rout, S. K. Bisoy and M. Sain, "Training Strategy of Fuzzy-Firefly Based ANN in Non-Linear Channel Equalization [J]," in IEEE Access, 2022,10(7): 51229–51241.
27. Li Chunguang, Liao Xiaofeng, Wu Zhongfu, et al. Adaptive Complex Channel Equalization Technology Based on a Hybrid Learning Algorithm [J]. Journal of Communications, 2001, (07): 38–43.
28. Cai Zejian, Ge Jianhua. Improved LMS semi blind adaptive channel equalization technology [J]. Television Technology, 2003,5 (11): 20–23.
29. Deng Xiaohong. Research on Adaptive Equalization Technology for Nonlinear Channels [D] Chengdu, Sichuan: Southwest Jiaotong University, 2004.
30. Sun Shouliang Communication Channel Equalization Theory and Application Technology [D]. Shandong: Shandong University, 2005.
31. Zhang Changchun, Yi Kechu. An adaptive equalization algorithm based on particle swarm optimization [J]. Electronic Technology, 2006,3 (11): 12–18.
32. Liu Jifeng, Du Shuoyi, Qiu Changxing. A Channel Equalization Algorithm Based on Gaussian Seidel Iteration [J]. Space Electronics Technology, 2006,69 (1): 73–83.
33. Long Ziqiang Research on Channel Equalization Technology in OFDM Systems for Wireless Communications [D]. Shantou, Guangdong: Shantou University, 2008.
34. Mi Bin, Yao Dalei, Lin Shouhui. Comparison of Learning Curves between Linear Equalizer and Decision Feedback Equalizer Algorithms [J]. Science and Engineering, 2007,11 (14): 3538–3540.

35. Chen Mingwu, Shao Chao. Adaptive DFE Channel Equalization Based on UKF [J]. Computer Development and Application, 2008,21 (12): 42–44.
36. Xu Shijun, Mo Lan. A New Variable Step Size Blind Constant Modulus Channel Equalization Algorithm [J]. Computer Simulation, 2008,25 (12): 139–141.
37. Li Jing Research on Information Channel Equalization Technology in Wireless Communication Systems [D] Dalian, Liaoning: Dalian University of Technology, 2011.
38. Zhao Zhijin, Zheng Xiaohua, Shen Lei. A second-order Volterra adaptive equalization algorithm based on lattice filtering [J]. Journal of Circuits and Systems, 2010,15 (06): 11–18.
39. Wang Xiangyu, Ge Wancheng. Research on Channel Equalization Algorithms Based on Turbo Iteration Technology [J]. Information Technology, 2011,35 (02): 24–26.
40. Wang Guifeng Research on Channel Equalization Based on Factor Graphs [D]. Dalian, Liaoning: Dalian University of Technology, 2012.
41. Shi Wei. Adaptive equalization technology based on the new quasi Newton method [J]. Journal of Beijing University of Posts and Telecommunications, 2012,35 (03): 83–86.
42. Wang Lin Research on Blind Equalization Technology for Wireless Communication Channels [D]. Hangzhou, Zhejiang: Hangzhou University of Electronic Science and Technology, 2013.
43. Song Yuwei, Yang Shouyi, Qi Lin. Nonlinear channel receiver equalization scheme based on Wiener model [J]. Television Technology, 2012,36 (05): 83–85.
44. Zhang Weiyi Research on Channel Estimation and Equalization Algorithms for SC-FDE Systems [D]. Chengdu, Sichuan: Southwest Jiaotong University, 2014.
45. Yang Yuan, Li Mingyang, Wang Xuhua. An improved sparse multipath channel equalization method [J]. Journal of Xi'an University of Electronic Science and Technology, 2014,41 (01): 158–163.
46. Xu Xiangyi, Xue Rui. Optimization Algorithm for Communication Channel Equalization Based on Particle Swarm Fitness Control [J]. Microelectronics and Computers, 2015,32 (07): 138–141.
47. Suri Research on Nonlinear Channel Identification and Equalization Technology Based on Hammerstein Model [D]. Zhengzhou, Henan: PLA University of Information Engineering, 2018.
48. Luo Xiaofang Research on equalization algorithms in visible light communication systems [D] Changsha, Hunan: Hunan University, 2018.
49. Li Liang. Nonlinear Channel Equalization Algorithm Based on Regenerative Kernel Hilbert Space [J]. Computer Engineering and Applications, 2016,52 (16): 105–1010.
50. Zhai Zheng'an, Liu Yang. Variable step size nonlinear channel equalization algorithm based on fractional lower order moments [J]. Computer Applications, 2016,36 (2): 11–14.
51. Wang He Channel estimation and equalization in atmospheric laser communication [D] Lanzhou, Gansu: Lanzhou University of Technology, 2016.
52. Chen Minhua, Li Yang, Zhang Wuxiong. Channel equalization algorithm based on convolutional neural networks [J]. Computer Application and Software, 2017,34 (09): 257–261.
53. Sun Danhua, Sun Liang, Wang Bin, et al α Nonlinear channel equalization algorithm based on kernel method under stable distributed noise [J]. Signal Processing, 2017,33 (02): 223–229.
54. Li Hao. Wireless Communication Channel Equalization Algorithm Based on Adaptive Bit Modulation [J]. Journal of Henan University of Engineering (Natural Science Edition), 2018,30 (03): 64–610.
55. Guo Hongchao Research on Channel Equalization Algorithms in Coherent Optical Communication Systems [D]. Beijing: Beijing University of Posts and Telecommunications, 20110.
56. Ma Zhenghua, Zhou Menglin, Chen Yang, et al. Channel equalization algorithm based on sparse source reconstruction [J]. Computer Engineering and Design, 2019,40 (09): 2489–2493.
57. Yang Rui, Sun Yunqiang, Du Xin, et al. Research on Channel Equalization Technology Based on NI USRP [J]. Foreign Electronic Measurement Technology, 2020,39 (09): 36–41.
58. Li Ru, Zhai Shuying, Bao Huiqin, et al. Simulation of Improved NLMS-OCF Algorithm in Channel Equalization [J]. Electronic Design Engineering, 2020,28 (16): 119–123.

59. Lu Xiaona. Research on Channel Equalization Based on LMS Algorithm [J]. Computer Knowledge and Technology, 2020,16 (17). 2022–216.
60. Bai Guo, Cheng Yufan, Tang Wanbin. Research on Single Carrier Frequency Domain Equalization Algorithm Based on Deep Learning [J]. Signal Processing, 2021,37 (06): 922–931.
61. Chen Bo, Liu Junjie, Zhang Shuo, et al. Channel Equalization Method for Coherent Optical Filter Bank Multicarrier Communication System Based on CD net [J]. Optical Communication Technology, 2022,46 (05): 70–76.
62. Wang Zaipeng Research on Low Complexity Channel Equalization Algorithms in Coherent Optical Communication Systems [D]. Qingdao, Shandong: Qingdao University, 2021.
63. Wang Chao Research on the Theory of Complex Gaussian Process Regression and Its Application in Channel Equalization [D]. Lanzhou, Gansu: Lanzhou University, 2022.
64. Ma Lina, Fang Xi, Zhang Lei. PDM O-OFDM/OQAM channel equalization method based on k-means [J]. Journal of Beijing Institute of Electronic Science and Technology, 2022,30 (03): 111–118.
65. Xie Duo. Research on Channel Equalization Algorithms Based on Shortwave Communication [J]. Changjiang Information Communication, 2022, 35 (04): 23–25.
66. Chen Yugeng Research on atmospheric laser communication channel equalization technology based on LSTM neural network [D] Changchun, Jilin: Changchun University of Science and Technology, 2022.
67. Menda Research on Channel Equalization and Coding Technology in Visible Light Communication Systems [D]. Changchun, Jilin: Jilin University, 2022.

Error-Correction Coding

18

Contents

Introduction 582
Research Progress in Error-Correction Codes 582
Applying Error-Correction Codes to Optical Wireless Communication 585
Error-Correction Coding Simulations in an Optical Wireless Communication System 587
- Error-Correction Control Based on RS Codes 588
- Error Control Based on LDPC Codes 591
- Error Control Based on Turbo Codes 591
- Error Control Based on Fountain Codes 591
- Error Control Based on Polarization Codes 596

Experimental Measurements 602
- RS Codes 603
- Turbo Codes 604
- LDPC Codes 605

Summary 605
References 611

Abstract

The channel environment of the optical wireless communication is complicated, susceptible to various natural phenomena such as rain, snow, and fog. The channel code is usually used to correct the error caused by the channel noise. Common channel codes are Reed–Solomon (RS) code, turbo code, low-density parity-check (LDPC) code, polarization code, and other types. The development of the error-correction code was combed, the research progress of the error-correction code for the optical wireless communication was summarized, and the comparative analysis was made, which introduced the progress of Xi'an University of Technology in this field of theory and experimentation. RS code, turbo code, and LDPC code were comparable to the error correction characteristics in rain, snow, and fog environments. Finally, the future possible research directions and challenges were pointed out, providing references for the future development and research in the relevant fields.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_18

Keywords

Optical wireless communication · Error-correction coding · Bit-error rate

Introduction

The optical wireless communication channel environment is complex and vulnerable to various natural phenomena, such as rain, snow, and fog. Channel coding methods are typically used to correct error codes caused by channel noise. Channel coding includes Reed–Solomon code (RS), turbo code, low-density parity-check code (LDPC), and polarization code.

This chapter presents the development of error-correction codes, summarizes the research progress on the error-correction codes used in optical wireless communication, and presents a comparative analysis. It introduces the theoretical and experimental research progress of Xi'an University of Technology in this field. In addition, it compares and analyzes the error-correction characteristics of RS, turbo, and LDPC codes in rain, snow, and fog. Finally, it lists possible future research directions and challenges in this field, which provides a valuable reference for future development and research in related fields.

In an era of rapid information technology development, cloud computing, artificial intelligence, and multimedia services continue to increase. Massive amounts of data are transmitted and stored in communication networks, and impose higher requirements on their bandwidth, speed, and delay. Optical communication, with wide bandwidth, large capacity, and high speeds, has been extensively developed.

Optical wireless communication transmits voices, images, text, etc., in a optical wireless channel. It has good application prospects in solving the "last mile" access, emergency rescue, disaster relief communication, etc. Natural phenomena in optical wireless channels adversely affect laser communication, such as broadening the communication pulse range, reducing the signal-to-noise ratio, and increasing the bit-error rate [1].

To suppress channel degradation, the anti-interference performance of signals must be improved through error-control coding (i.e., error-correction codes) and increasing the signal redundancy to meet the system index requirements. Introducing error-correction codes into optical wireless communication systems is another new practice in error-correction code applications. The research and development of coding technologies with good anti-interference performance and high speeds have become key topics in the field of optical wireless communication [2].

Research Progress in Error-Correction Codes

Shannon founded information theory in 1948 and published the pioneering document "Mathematical Theory of Communication," marking the birth of information and coding theory and the beginning of error-correction coding theory [3]. Since

then, many scholars have devoted themselves to researching error-correction coding. With improvements in coding methods and computing power, error-correction coding methods that have better error-correction performance and are closer to the Shannon limit have emerged one after another.

In 1950, Hamming proposed a linear block code, called the Hamming code, which could correct an independent error. The Hamming code can indicate the location of a data error and correct a single-bit error in seven bits [4]. In 1954, Muller proposed a structurally rich algebraic coding scheme [5]. In the same year, Reed of the Massachusetts Institute of Technology of the United States proposed an efficient decoding algorithm for the hierarchical structure of the algebraic coding scheme: a large-number logical decoding algorithm. This algebraic coding scheme is called the Reed–Muller (RM) code. The RM code is also a linear block code, but it can correct multiple errors, unlike the Hamming code [6]. It was later applied to Mars exploration.

In 1954, Golay studied the shortcomings of the Hamming code and proposed the Golay code, which can correct any three-bit errors [7], effectively improving the coding efficiency. The abovementioned error-correction codes are all linear block codes.

In 1955, Elias proposed convolutional codes, which are non-block codes. They are more suitable for forward error correction. In many practical applications, their performance is better than that of linear block codes, with a small amount of computation. This coding method makes better use of the correlation between various modules. The coding output at the current time depends on the input at the current time and the input at the previous time; thus, a higher coding gain can be obtained. The emergence of convolutional codes also greatly improved the development of wireless communication [8].

In 1957, Prange proposed the cyclic code, which improves on linear block codes. It not only has the characteristics of general linear block codes but also has the cyclic characteristics of any code word after a cyclic shift [9]. Cyclic codes have strong error detection and correction capabilities, can detect random and some burst errors [10], and the encoding and decoding equipment is relatively simple; however, their ability to correct multiple random errors is limited.

In 1959, Hocquenghem proposed the Bose–Chaudhuri–Hocquenghem (BCH) code, which is a cyclic code that can correct multiple random errors and compensate for the limited correction ability of the cyclic code proposed by Prange. It has been studied extensively [11].

In 1960, Reed and Solomon proposed the Reed–Solomon (RS) code that effectively corrects random-symbol and random-burst errors. This code is a further improvement on the BCH code with strong error-correction ability [12]. It is especially applicable to multi-ary modulation.

Cyclic codes add linear constraints and are the most mature type of codes. Most practical error-correction codes are cyclic (Table 18.1).

The coding efficiencies of several error-correcting coding schemes proposed by scholars were very close to the Shannon limit, which significantly promoted the development of error-correcting codes. The low-density parity-check (LDPC) code

Table 18.1 Research progress on error-correction codes

Year	Researcher	Research progress
1948	Shannon	Published *A Mathematical Theory of Communication*
1950	Hamming	Hamming code—Linear block code for correcting independent errors
1954	Reed and Muller Golay	RM code Golay code
1955	Elias	Convolutional code
1957	Prange	Cyclic code
1959	Hocquenghem	BCH code
1960	Reed and Solomon	RS code
1962	Gallager	LDPC code
1993	Berrou and Glavieux	Turbo code
2009	Arikan	Polar code
2006	Li	BCH-LDPC. Concatenated code with better performance
2007	Kang	RS-BCH. Concatenated code used to correct mixed random and burst errors
2009	Zhang	RS-Turbo. Concatenated code improves coding gain
2010	Yuan	RS-BCH. Concatenated code improves the system error-correction ability
2017	Dong	Research on irregular QC-LDPC code construction

proposed by Gallager in 1962 is a linear error-correction code based on a sparse matrix, with performance very close to the Shannon limit. The advantages of this code are more significant when a code group is extremely long.

At that time, computers were not able to process lengthy code groups; thus, the LDPC code did not receive scholars' attention [13, 14]. Later, with the enhancement of computing ability, it received renewed attention and is now widely used in many fields, such as mobile and optical fiber communication.

In 1993, Berrou and Glavieux proposed the turbo channel coding scheme with a coding efficiency approaching the Shannon limit [15]. The proposed turbo code broke away from the traditional channel coding method, integrated a convolutional encoder and interleaver, effectively reduced the correlation between codes, and rendered the coding sequence highly random. The decoding algorithm used for this code is extremely complex.

The encoding efficiencies of the LDPC and turbo codes are very close to the Shannon limit; however, their decoding delay is high. Therefore, they are more suitable for communication systems with low real-time requirements. Arikan et al. proposed the polar code in 2009; it is a channel coding, error-correcting scheme with relatively low coding and decoding complexities, which has been proven to theoretically approach the Shannon limit. It has been proved theoretically to reach the capacity of any binary input discrete memoryless symmetric channel [16, 17].

We have discussed the emergence, advantages, and disadvantages of various error-correction codes. During the development of error-correction codes, scholars have improved various coding methods and concatenated different error-correction codes, resulting in different combinations of concatenated codes, and thereby improving the coding performance.

In 2006, Li et al. concatenated the BCH and LDPC codes. The error rates of the concatenated and LDPC codes were similar, but the concatenated code's signal-to-noise ratio performance was better [18]. In 2007, Kang et al. used concatenated RS and BCH codes to correct random and burst errors, which effectively reduced the error rate in a Poisson channel [19].

In 2009, Zhang et al. found that the performance of concatenated RS and turbo codes as an error-correction code was at least 1 dB higher than that of the concatenated RS and convolutional codes. A bit-error rate of 10^{-6} resulted in a 2-dB coding gain [20].

In 2010, Yuan et al. proved through simulation that when the bit-error rate was the same, concatenated RS and BCH codes achieved a greater coding gain, compared to using the RS code alone, and the system error-correction performance was greatly improved [21]. In 2017, Dong et al. investigated the optimal allocation of sub-channel rates and the construction of irregular quasi-cyclic low-density parity-check (QC-LDPC) codes, focusing on the optimization of multilevel coding and modulation schemes [22].

The above results show that concatenated codes can effectively improve the error-correction performance and coding efficiency, and combinations of different coding methods are better than single codes. The emergence of various concatenated codes has further promoted the development of error-correction coding. The development of error-correction coding is presented in Table 13.1.

Applying Error-Correction Codes to Optical Wireless Communication

Optical wireless communication uses an atmospheric channel, and the communication link is extremely vulnerable to atmospheric effects. A poor channel environment reduces the communication quality. Channel coding suppresses the interference of an optical signal in a channel. Channel coding can improve the signal anti-interference ability and the system's communication performance.

Forward error-correction codes are generally applied to optical wireless communication systems. In 1993, Ohtsuki et al. applied the RS code to an optical wireless communication system and analyzed the results [23]. In 2003, Josephm et al. combined turbo codes with an optical wireless communication system and conducted theoretical calculations and simulations for this scheme [24]. In 2005, Jaime et al. analyzed the atmospheric channel capacity of an optical wireless communication system from the perspective of information theory, introduced an LDPC code to suppress the channel noise, and obtained a large net coding gain [25].

The application of RS, turbo, and LDPC codes to optical wireless communication systems can improve the system performance and increase the coding gain. Scholars have compared several coding methods from different perspectives. In 2008, He et al. deduced the average bit-error rate (BER) of an optical wireless communication system when the channel information was known and compared the system performance of two schemes using RS and turbo coding. The results showed that the turbo coding scheme effectively reduced the required reception power.

In weak turbulence, the BER of turbo codes slowly deteriorates with an increase in turbulence intensity, making them more suitable for optical wireless communication systems than the RS code. Later, scholars began applying concatenated codes to optical wireless communication systems [26].

In 2010, Wen et al. studied channel coding in satellite communications. The fountain code and nonbinary LDPC codes were introduced, including their encoding and decoding algorithms. Cascade and nonbinary LDPC fountain codes were proposed based on characteristics similar to those of sparse graphic codes. Simulation results demonstrated the good performance of the cascade codes [27].

In 2011, Jiang et al. optimized the performance of interleaved RS codes. The optimal coding efficiency was calculated when the interleaving depth was fixed. Simulation results showed that interleaved RS codes significantly improved the performance of communication systems in fading channels. They have the maximum anti-interference ability when the coding efficiency is 0.5 [28].

The excellent performance of polarization codes has attracted the attention of many researchers. In 2013, Habif et al. combined a polarization code with pulse-position modulation and adopted direct detection to achieve a photon information efficiency of 4.8 bits per received photon in a weakly coherent state [29]. In 2015, Wu et al. applied polarization codes to an optical communication system and studied the transmission performance of a 7% redundant polarity code in a coherent ultra-dense wavelength-division multiplexing (WDM) optical access network scheme. They obtained a performance gain of 1.4 dB when the bit-error rate was 10^{-3} [30].

In 2018, Fang et al. analyzed the performance of polar codes in turbulent atmospheric channels [31]. The results showed that the polar code could effectively suppress the signal fading caused by atmospheric turbulence in medium- and strong-turbulence channels, and its coding complexity was low.

In 2019, Iqbal et al. proposed a coherent-optical polarity-coding transmission system with orthogonal amplitude modulation and an adjustable data rate. An implicit hole–hole polar-coded modulation system was designed, based on a many-to-one constellation-shaping method. The scheme was experimentally evaluated in a WDM system, and a new hole polarity code was proposed. Compared with the traditional hole polarity code, the shaping gain was 1.2 dB, and the maximum communication distance was increased by 200–400 km [32].

To improve the transmission efficiency in weak atmospheric turbulence channels, Zhang et al. designed an adaptive-rate polarization code in 2021. Under different turbulence intensities, when the frame error rate was 10^{-8}, the adaptive-rate polarization code obtained a performance gain of 1.7–2.3 dB [33]. The above conclusions were mostly based on theoretical calculations and simulations.

In 2019, Ao et al. used the Raptor10 code to combine optical wireless and radio frequency (RF) channels. Optical wireless communication was demonstrated on a 1.87-km urban link, and the Raptor code was applied to the actual channel coding, which proved that it could maximize the link utilization [34].

A multiple-input multiple-output (MIMO) transmission and reception method was proposed to suppress the impact of atmospheric turbulence on communication systems. In 2021, Wen et al. proposed a MIMO polarization-coding scheme suitable for optical wireless communication to address the problem of communication links being vulnerable to atmospheric turbulence. It maintained low complexity while achieving better system performance [35].

Scholars have also improved the methods for decoding polarization codes. In 2021, Wenhao et al. proposed a segmented polarization code–cyclic redundancy check–serial cancellation list decoding method based on relaxed polar codes. This method significantly improved the error rate of free-space optical (FSO) communication systems and the signal-to-noise ratio [36].

In the same year, Lu et al. analyzed channel combination and decomposition in channel polarization, further studied the polarization code encoding and decoding process, and discussed the polarization calculation method in different channels and the impact of coding parameters on the bit-error rate [37].

The optimization and improvement of other coding schemes have continued. Research on LDPC codes in optical wireless communication systems is becoming increasingly mature. In 2021, Cao et al. proposed a dynamically adjusted log-likelihood ratio (LLR) as the input of an LDPC decoder to obtain better communication performance. Compared to the traditional LLR input, the simulation results showed that the LDPC code based on the dynamically adjusted LLR had strong turbulence resistance and better error performance under different atmospheric turbulence conditions [38].

In 2021, Sonali et al. proposed an analysis model verified by simulation to evaluate the performance of LDPC codes in optical wireless communication links under various channel conditions. They analyzed and formulated a backpropagation (BP) algorithm using LDPC codes in fast-fading and spatial-multiplexing optical wireless communication channels. The maximum coding gains were 13.95 dB and 19.83 dB, respectively [39].

Some applications of error-correction codes in optical wireless communication are listed in Table 18.2.

Error-Correction Coding Simulations in an Optical Wireless Communication System

The performance of an optical wireless communication system is affected by several factors. Alignment errors can cause spot drift, reduce the signal-to-noise ratio, and cause the optical signal to randomly fluctuate. Atmospheric absorption can cause the power to attenuate and reduce the average signal-to-noise ratio. Scattering can cause optical power attenuation and optical pulse broadening, resulting in intersymbol

Table 18.2 Applications and research progress of error-correction codes in optical wireless communication

Year	Researcher	Research progress
1993	Ohtsuki	RS code used in optical wireless communication system
2003	Josephm	Used turbo codes in optical wireless communication systems
2005	Jaime	Introduced LDPC code to suppress channel noise
2008	He	Turbo code requires lower reception power
2010	Wen	Performance of cascade fountain code and binary LDPC code is better
2011	Jiang	Interleaved RS code has the strongest anti-interference ability when the coding efficiency is one-half
2013	Habif	Combined a polarization code with pulse-position modulation
2015	Wu	Polarization code used in an optical communication system
2019	Iqbal	Design of an implicit drilling polarity-code modulation system
2019	Ao	Used Raptor10 code in a combined FSO and RF channel
2019	Zhang	Adaptive-rate polarization codes for weak-turbulence channels
2021	Wen	Polarization code used in MIMO system with FSO communication PC-CRC-SCL decoding method for an FSO communication system
2021	Lu	Influence of polarization calculation method and coding parameters on the bit-error rate in different channels
2021	Cao	Performance of LDPC code based on a dynamically adjusted log-likelihood ratio (LLR) is better
2021	Sonali	Analyzed and developed a backpropagation algorithm for LDPC code

crosstalk and a lower signal-to-noise ratio. Atmospheric turbulence can cause light intensity fluctuation, beam drift, angle-of-arrival fluctuation, and beam expansion. It also distorts the optical signal wavefront, reduces the average signal-to-noise ratio, and generates random fluctuations in the received signal.

This chapter only analyzes the impact on the BER performance. Since 2002, Xi'an University of Technology has studied error-control coding in optical wireless communication systems. RS, turbo, LDPC, fountain, and polarization codes have been used in optical wireless communication systems, and the performance of various codes in optical wireless communication systems has been simulated and experimentally studied.

Error-Correction Control Based on RS Codes

An RS code groups the original information and adds a certain number of supervision symbols to each information group. Every nine symbols are divided into a group with six bits of supervision symbols added. The RS decoder adopts a pipeline decoding algorithm. We simulated the performance of the RS (30, 10) code used in optical wireless communication systems under various weather conditions [40] (Figs. 18.1, 18.2, 18.3, and 18.4).

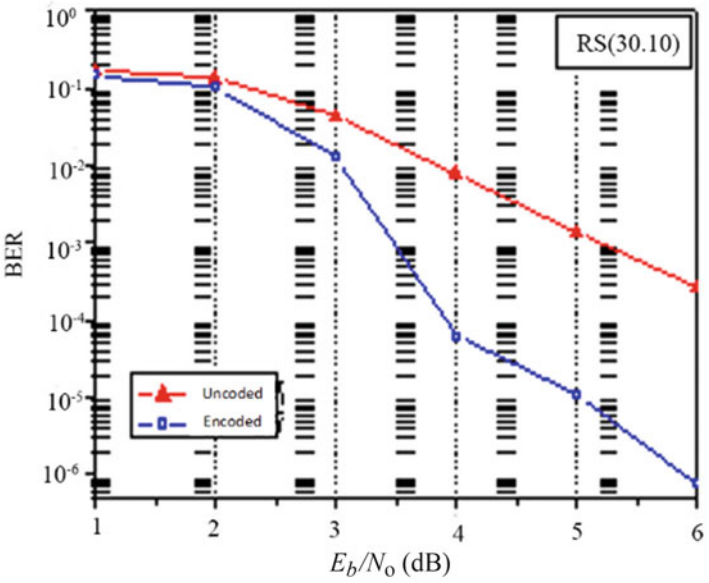


Fig. 18.1 Error rate before and after error correction with RS (30, 10) code on sunny days [40]

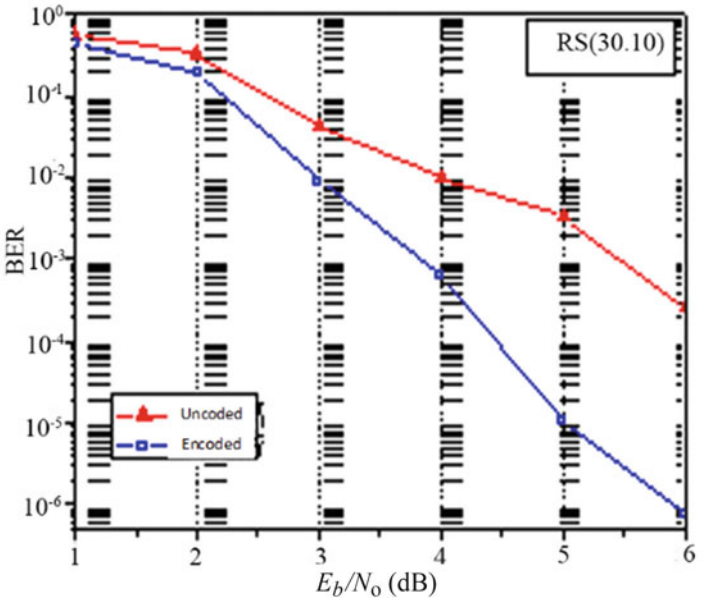


Fig. 18.2 Error rate before and after error correction with RS (30, 10) code on cloudy days [40]

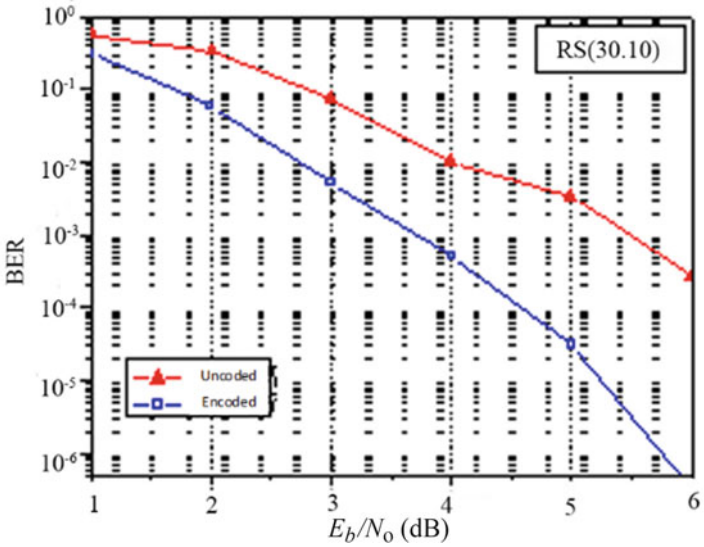


Fig. 18.3 Error rate before and after error correction with RS (30, 10) code on rainy days [40]

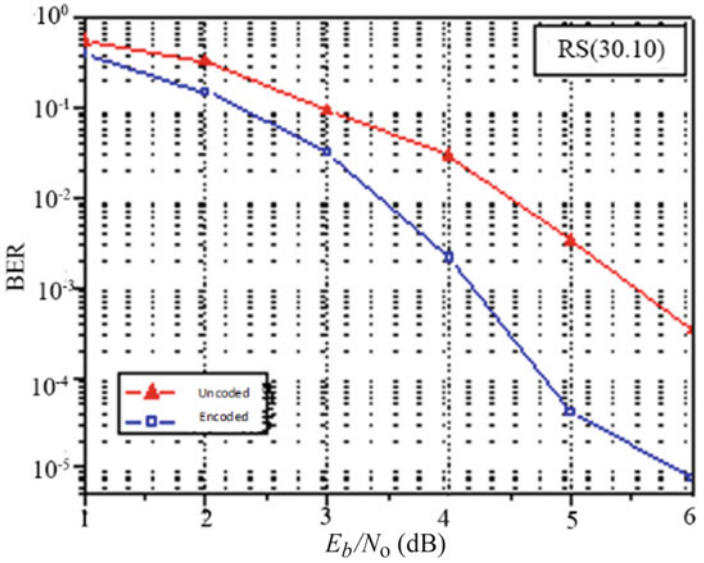


Fig. 18.4 Error rate before and after error correction with RS (30, 10) code on snowy days [40]

Error Control Based on LDPC Codes

The LDPC code is a linear error-correction code with excellent performance that comes very close to the Shannon limit. Figures 18.5, 18.6, 18.7, and 18.8 show the error-correction performance of LDPC codes under different weather conditions [41].

Error Control Based on Turbo Codes

Figures 18.9, 18.10, 18.11, and 18.12 compare the performance of the bit-error rate before and after turbo code error correction under four different weather conditions: sunny, cloudy, rainy, and foggy [43].

Figures 18.9, 18.10, 18.11, and 18.12 show that the error rate with the turbo codes in different weather conditions improved to different degrees. Although the system was complicated by error-correction coding, its performance was greatly improved [43].

Error Control Based on Fountain Codes

Fountain codes are error-correction codes that generate multiple coded packets from one original data packet. All the original data packets can be recovered by receiving any coded packet. Luby transform (LT) codes and Raptor codes are specific implementations of fountain code [45].

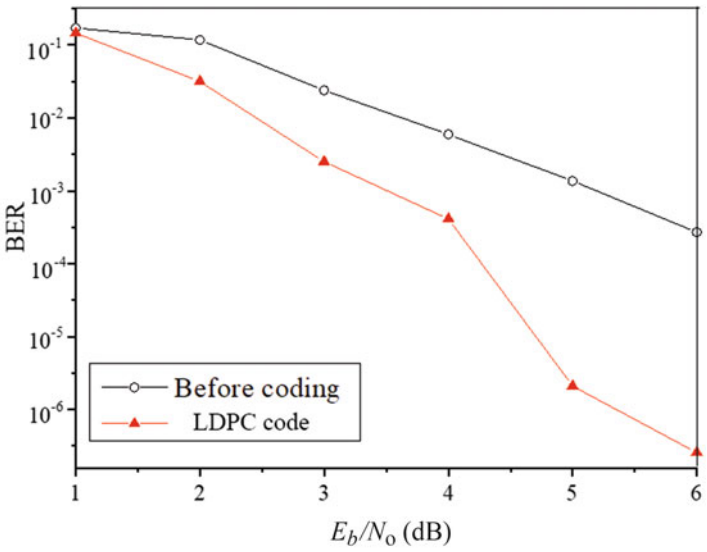


Fig. 18.5 Error rate before and after LDPC code error correction on sunny days [42]

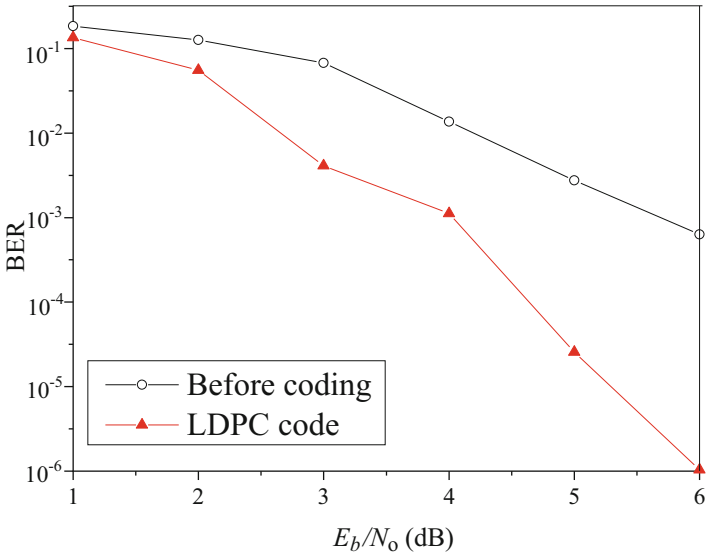


Fig. 18.6 Error rate before and after LDPC code error correction on cloudy days [42]

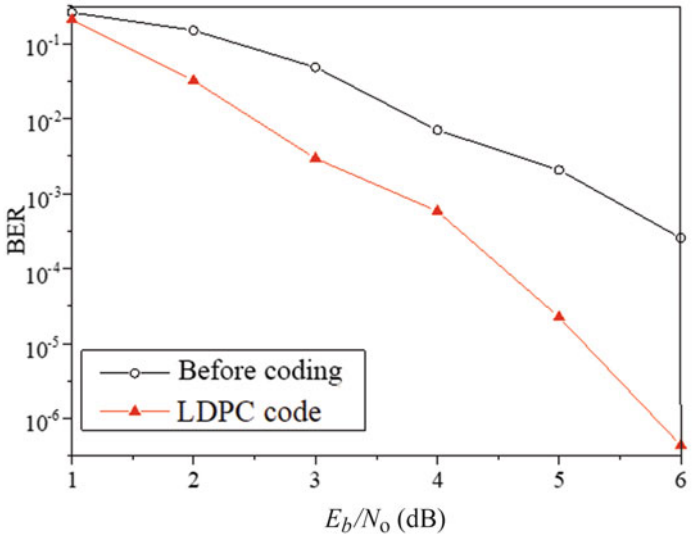


Fig. 18.7 Error rate before and after LDPC code error correction on rainy days [42]

LT Codes

The decoding success probability of LT codes is closely related to the degree distribution. Luby provides two practical degree distributions: ideal soliton and robust soliton. Because the number of verification units with an ideal soliton

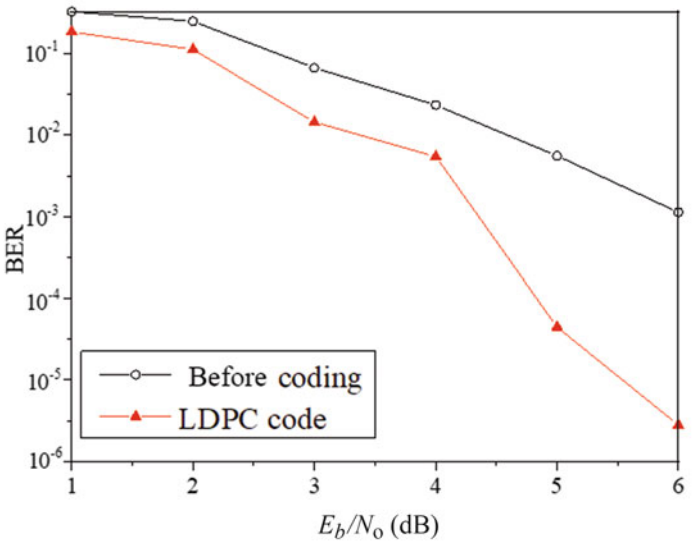


Fig. 18.8 Error rate before and after LDPC code error correction on snowy days [42]

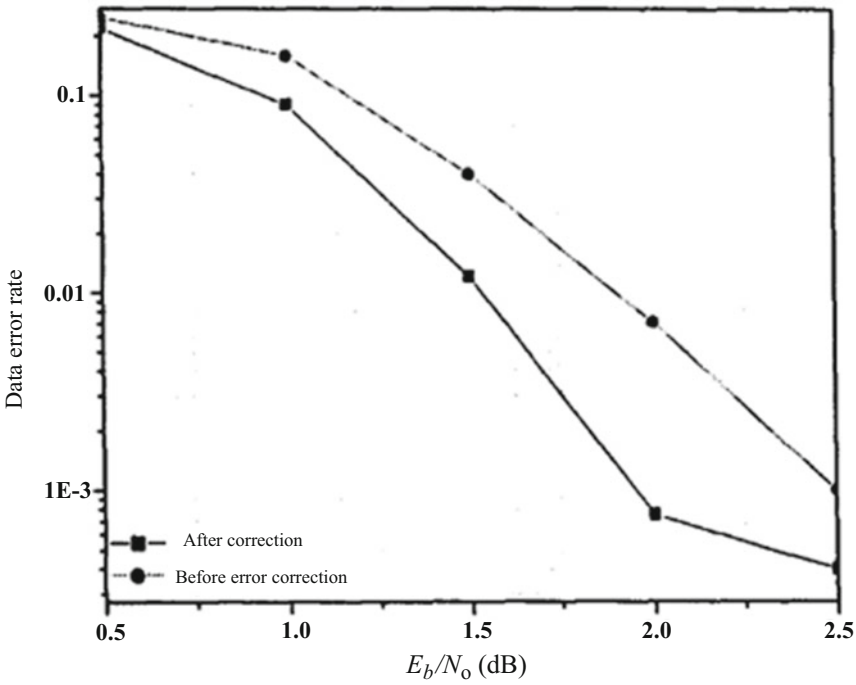


Fig. 18.9 Error rate before and after error correction using turbo codes on sunny days [44]

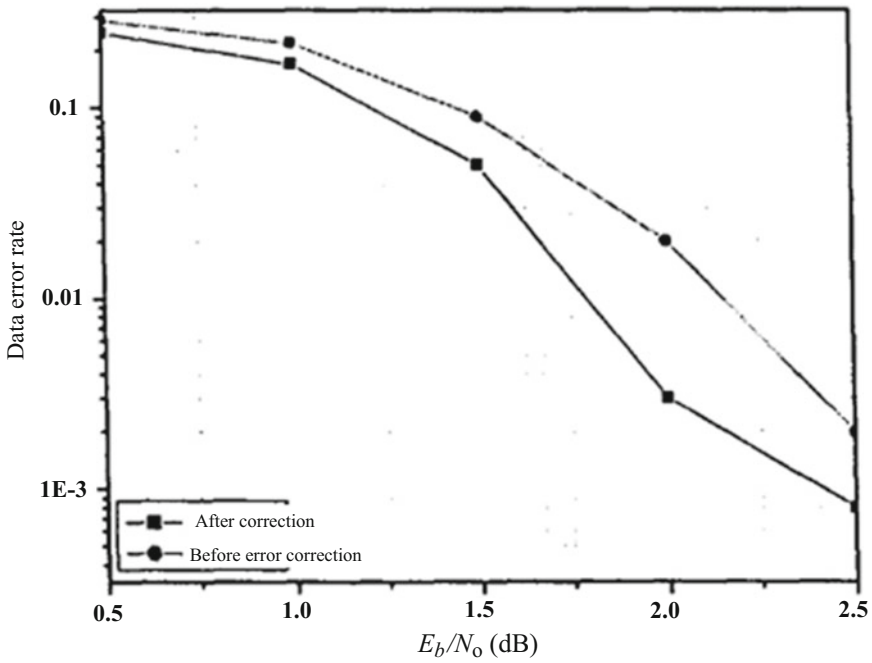


Fig. 18.10 Error rate before and after error correction using turbo codes on cloudy days [44]

distribution degree of one only accounts for the total number of verification units, the number is small and the fluctuation is large. Sometimes, there are no nodes with a degree of one, and some original packets have no verification nodes connected to them at all. As a result, the original information cannot be recovered correctly from the received information. This chapter adopts an improved robust soliton distribution that adds two parameters, which can achieve better results than the ideal soliton distribution [45].

Figure 18.13 compares the performance of LT codes with two different degree distributions: ideal and robust solitons. The abscissa represents the coding overhead, defined as $\varepsilon = \frac{k}{n} - 1$, and the ordinate represents the decoding failure probability, defined as $k = 155$, $c = 0.8$, $\delta = 0.05$. As shown in Fig. 18.14, when $n = 1.6k$, the decoding failure rate under the ideal soliton distribution is approximately 0.15, while the failure rate using the robust soliton distribution is 0.1; thus, the decoding failure probability is significantly reduced [45].

Raptor Codes

After decoding the LT code, $k' < k$ original input symbols are recovered, and the recovery failure rate is defined as $\eta = (k - k')/k$. As long as η is less than the error-correction ability of the RS code, the code can recover incorrect symbols. Figure 18.15 compares the performance of the LT and Raptor codes under the same conditions, where $k = 155$, $c = 0.8$ and $\delta = 0.05$ [45].

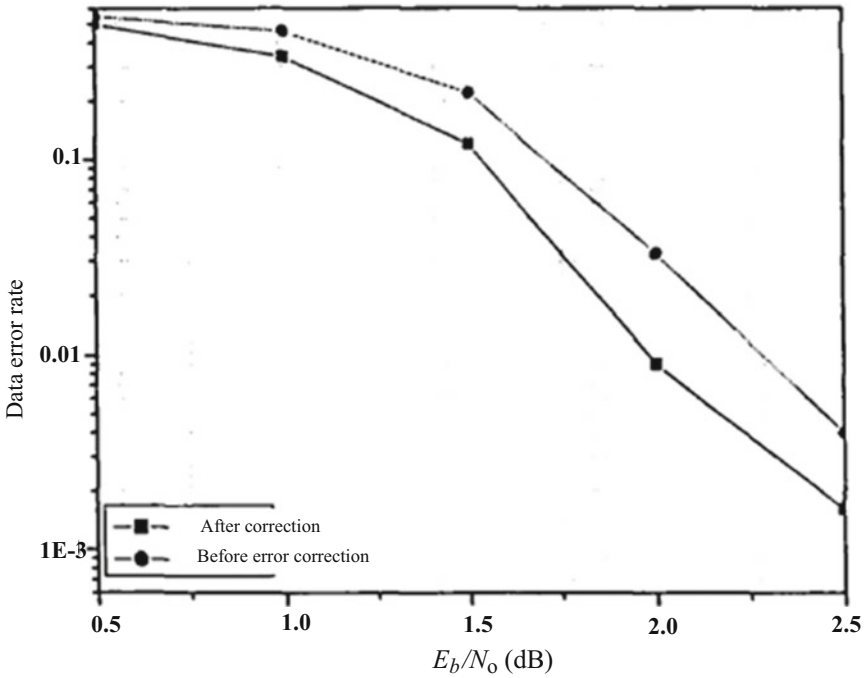


Fig. 18.11 Error rate before and after error correction with turbo codes on rainy days [44]

After adopting the Raptor code, the RS code can correct errors that cannot be recovered by the LT code, which significantly reduces the error rate. When receiving a coded packet with $n = 1.6k$, the decoding failure probability is approximately 0.12 when using the LT coding method, and approximately 0.08 when using the Raptor code. Therefore, the Raptor code is greatly optimized [45].

The higher the value of the fountain code, the lower the probability of decoding failure, which is suitable for transmission of larger data volumes. Figure 18.16 compares the decoding failure probability $c = 0.8$, $\delta = 0.05$ [45] under the same reception redundancy as $k = 100$, $k = 1000$.

As shown in Fig. 18.16, the decoding failure probability when $k = 1000$ is lower than that when $k = 100$. When the received coding packet $n = k$ is received, the decoding failure probability of the longer data packet is 0.98 and the decoding failure probability of the shorter data packet is 1. When the data are grouped as $n \geq 1.1k$, the decoding failure rate decreases rapidly. When encoding the packet $n = 1.55k$, the longer data packet is recovered, whereas the decoding failure probability of the shorter data packet is approximately 0.1 [45].

When the distribution is certain, the more the original input data are grouped, the more the BER changes with increasing overhead, which has a significant advantage compared to small data volumes. When the received data packet n is

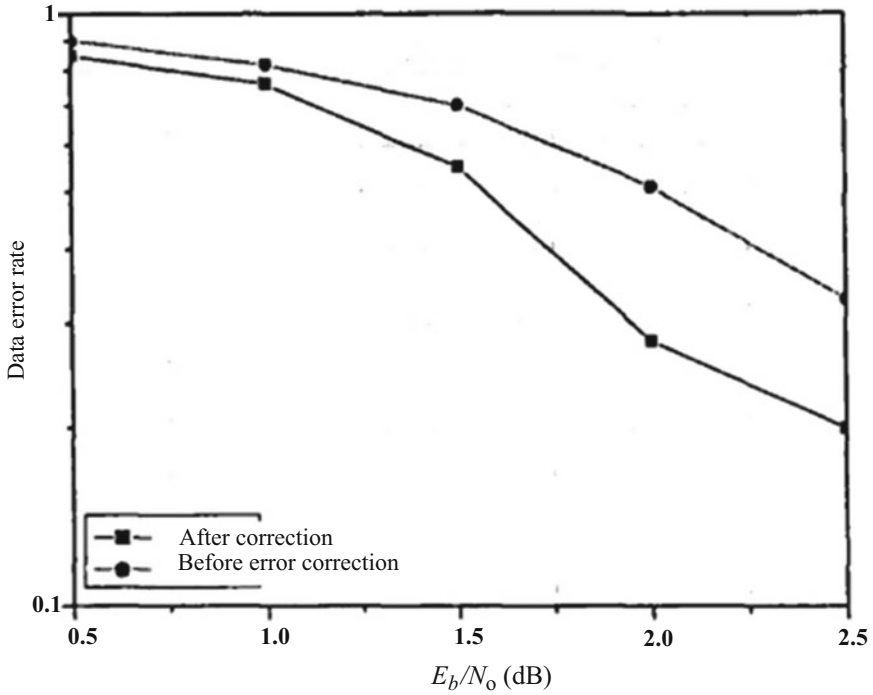


Fig. 18.12 Error rate before and after error correction with turbo codes on foggy days [44]

slightly larger than $k(n > 1.1k)$, the probability of a decoding failure decreases rapidly. However, when n is large, the change of bit-error rate is relatively slow. With an increase in the received coded packets, the decoding success probability also increases and is related to the number of original data packets [45].

Error Control Based on Polarization Codes

The polarization code has low coding complexity and is the only error-correction code that can be proven to approach the Shannon limit. This section analyzes the performance of the serial-cancellation (SC) decoding algorithm and the impact of pulse-position modulation (PPM) on the performance of polarization codes in an atmospheric laser channel. Experimental results demonstrated the error-correction performance of the polarization code under subcarrier modulation [46, 47].

Polarization codes can reach the channel capacity and fulfill the requirements of 5G channel coding schemes. Therefore, achieving an efficient coding design with limited code length, and developing high-efficiency and high-spectral-efficiency

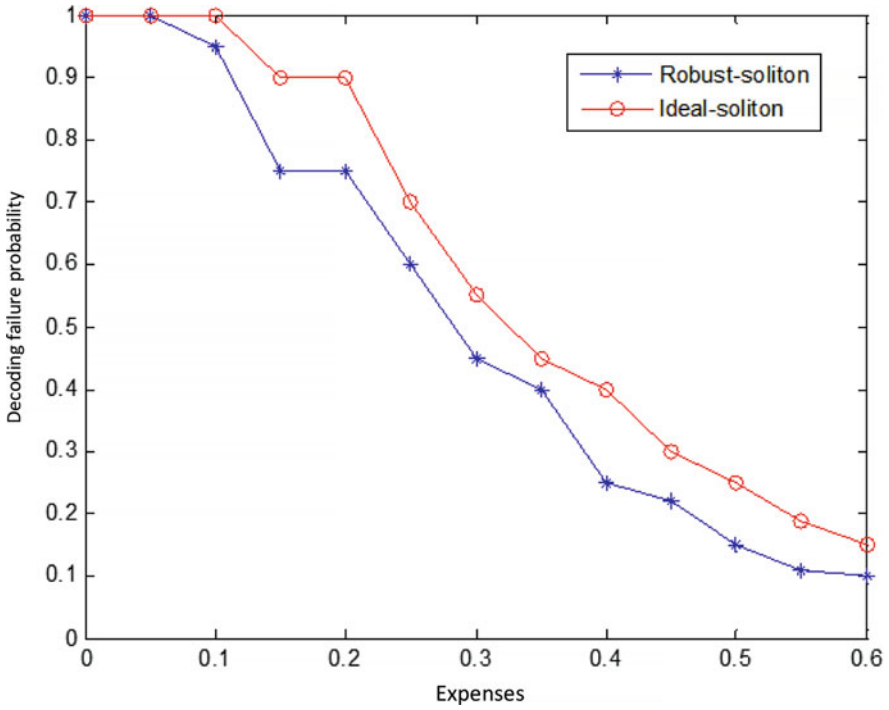


Fig. 18.13 Comparison of LT code performance under two degree distributions [45]

data-transmission technology combined with MIMO and non-orthogonal multiple access have become important research directions for improving future 6G wireless-transmission performance [48].

SC Decoding Algorithm Performance Simulation

Figure 18.16 shows the decoding simulation results of the serial-cancellation (SC) polarization code algorithm in a weak-turbulence channel constructed using a Gaussian approximation. It shows the bit-error rate of the polarization code in a weak turbulence channel when the signal-to-noise ratio is large (10^{-6}). Figure 18.16a compares the bit-error rate performance when $R = 0.5$ and $N = 512$, 1024 and 2048.

Under the same bit rate, the performance of the polarization code tends to improve with an increase in code length. The ratio of the 1024 code length to the 512 code length has a 0.8-dB coding gain, while the ratio of the 2048 code length to the 1024 code length has a 1.5-dB coding gain. It can be seen that increasing the code length can improve the coding gain of the polarization code. However, it increases the cost and delay of the encoder and decoder; the progressive complexity of coding and decoding is $O(N \log N)$.

Figure 18.16b shows the impact of different code rates on the polarization code when the code length is the same. It can be seen from the figure that the code rate has

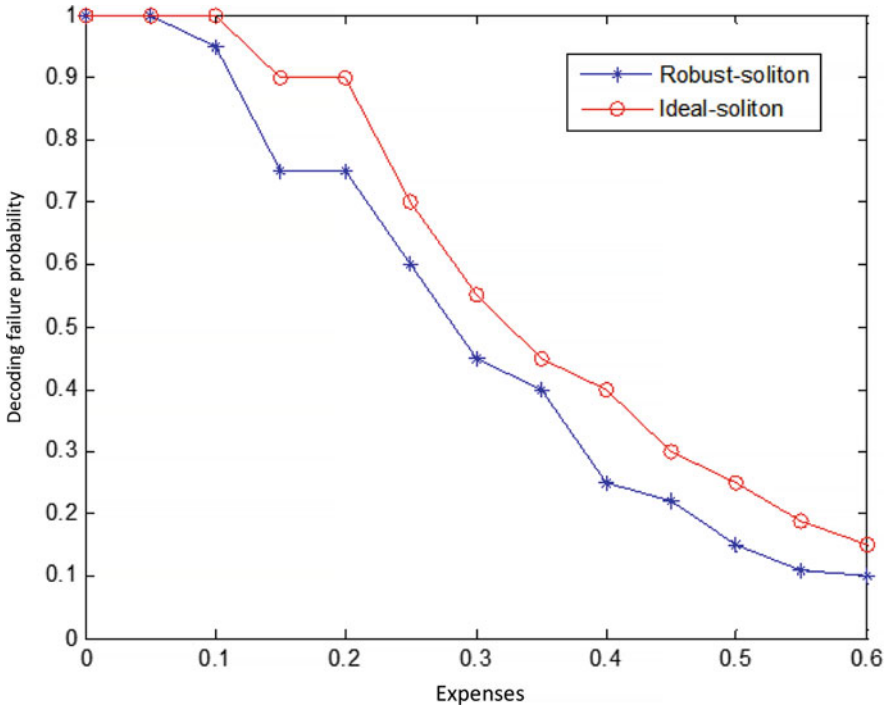


Fig. 18.14 Comparison of LT code and Raptor code performance [45]

a very large impact on the polarization code. This is because, with an increase in the bit rate, more channels are not fully polarized for information transmission. The impact of noise on this part of the channel is more evident with an increase in the bit rate, which increases the bit-error rate and reduces the decoding performance [46, 47].

We consider the impact of weak turbulent noise on the communication system when the code length is $N = 1024$ and the code rate is $R = 0.5$. Figure 18.17 shows the BER simulation curve of the SC algorithm decoding under different turbulence intensities. It can be observed that the system performance is reduced by atmospheric turbulence, and the reduction is more evident with an increase in turbulence intensity. For the same code rate and code length, the stronger the turbulence, the fewer channels are fully polarized, and the greater the bit-error rate [47].

Polarization Code Performance Under PPM Modulation

The bit-error rate performance simulation diagram of binary on-off keying (OOK) modulation and 2-PPM modulation is shown in Fig. 18.18, and the bit rate is 0.5. As shown in Fig. 18.18, the performance of the 2-PPM modulation is significantly improved, compared with that of OOK modulation. Under the same code-length and

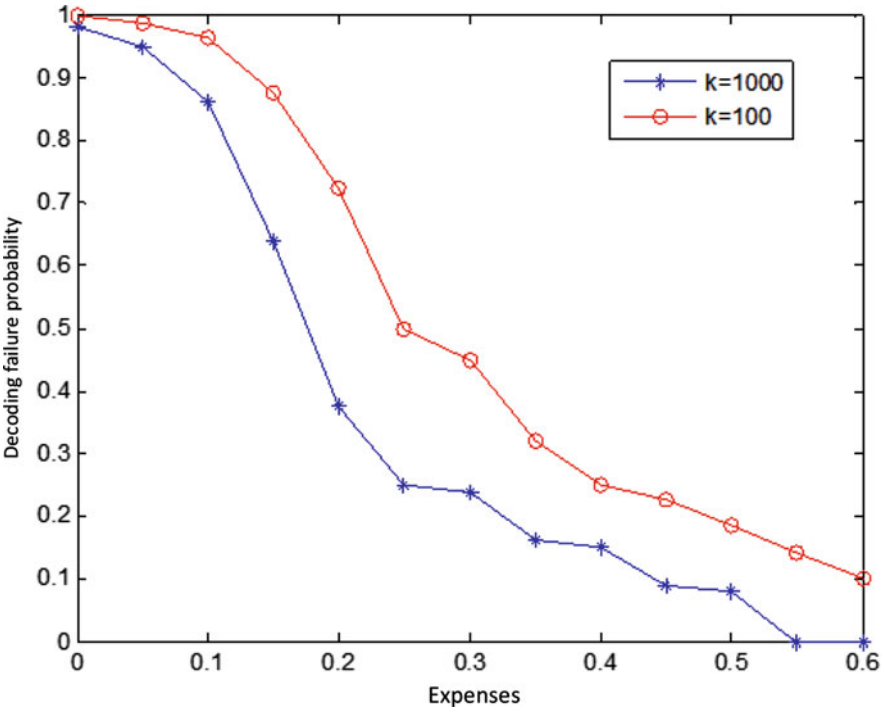


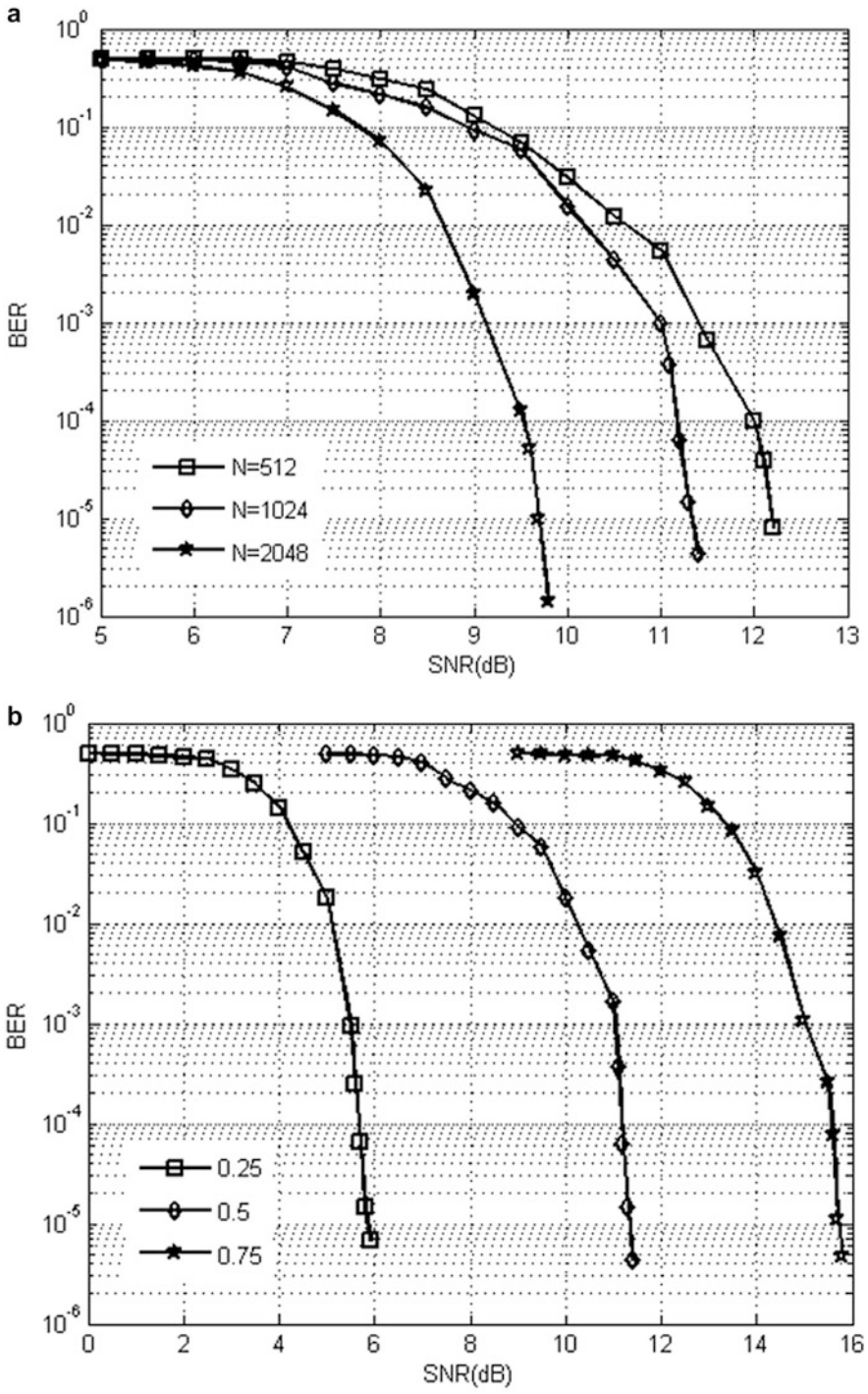
Fig. 18.15 Decoding failure rate at $k = 100$, $k = 1000$ [45]

code-rate parameters, 2-PPM modulation can obtain a 3.5 dB–3.8 dB coding gain compared to OOK modulation [47].

Figure 18.19 shows the performance of polarization codes in an atmospheric turbulence channel with an increase in M-ary PPM. As can be seen from the curve in Fig. 18.19, when the number of slots is small, the bit-error rate decreases with an increase in M, and the signal-to-noise ratio gain also increases.

However, as the modulation order M continues to increase, the modulation performance decreases; for example, the performance of 16-PPM is worse than that of OOK. This is because, the error probability of each frame demodulation increases with an increase in the modulation order M, which also increases the influence of turbulence and Gaussian noise, and thereby reduces its demodulation and decoding performance.

However, the average laser power required by the transmitter decreases with an increase in the number of time slots. The use of higher-order PPM results in a higher power utilization rate, which can improve the effective communication distance for optical wireless communication. Therefore, a reasonable compromise must be made in actual application scenarios [47, 49].


a
BER
SNR(dB)
N=512
N=1024
N=2048
b
BER
SNR(dB)
0.25
0.5
0.75

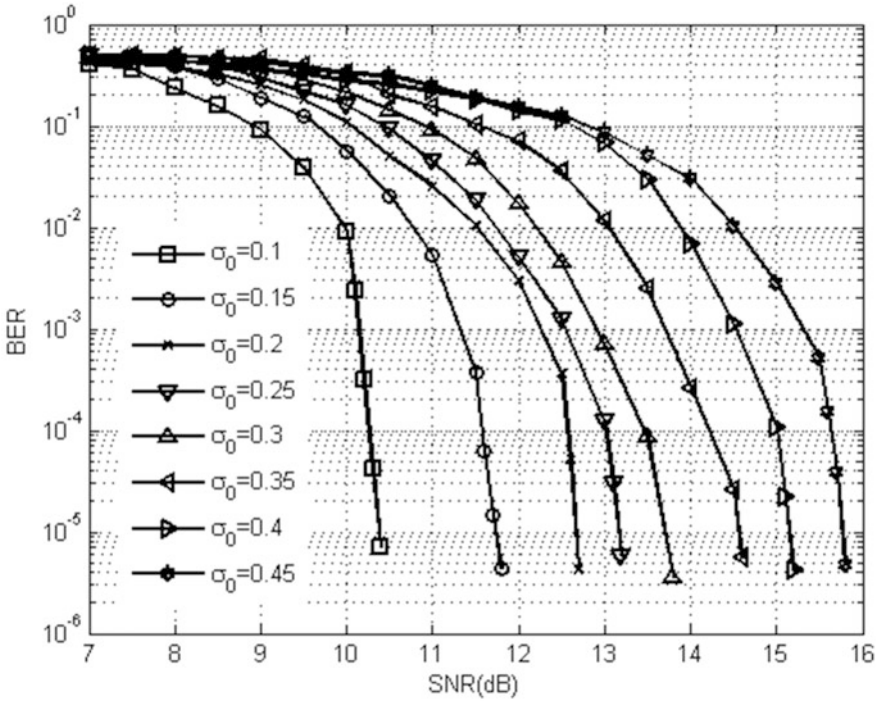


Fig. 18.17 Effect of different turbulence intensities on the bit-error rate [47]

Simulation and Experiment of Polarization Code Performance Under Subcarrier-Modulation Technology

The SC decoding algorithm was used in the experiment and simulation in this section, with a code length of 512 and code rate of 0.5. Figure 18.20 shows the BER simulation curve of uncoded subcarrier binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) modulations under different turbulence intensity channels ($\sigma_0 = 0.1$, $\sigma_0 = 0.2$, and $\sigma_0 = 0.3$) and different signal-to-noise ratios (SNRs) [47].

Figures 18.21 and 18.22 show the bit-error rates of subcarrier BPSK and QPSK modulations in different turbulence intensity channels, respectively. Comparing Figs. 18.21 and 18.22, it can be seen that the BPSK system using polarization coding is superior to QPSK and obtained a 4.8-dB coding gain [42].

Table 18.3 shows the bit-error rate characteristics of coded and uncoded data on sunny, cloudy, and rainy days. Using a polarization code can improve the bit-error

Fig. 18.16 Polarization code performance curve under a weak-turbulence channel [42]. (**a**) Constant code rate R = 0.5; different code lengths N = 512, 1024, and 2048 [47]. (**b**) Constant code length N = 1024; different code rates R = 0.25, 0.5, and 0.75 [47]

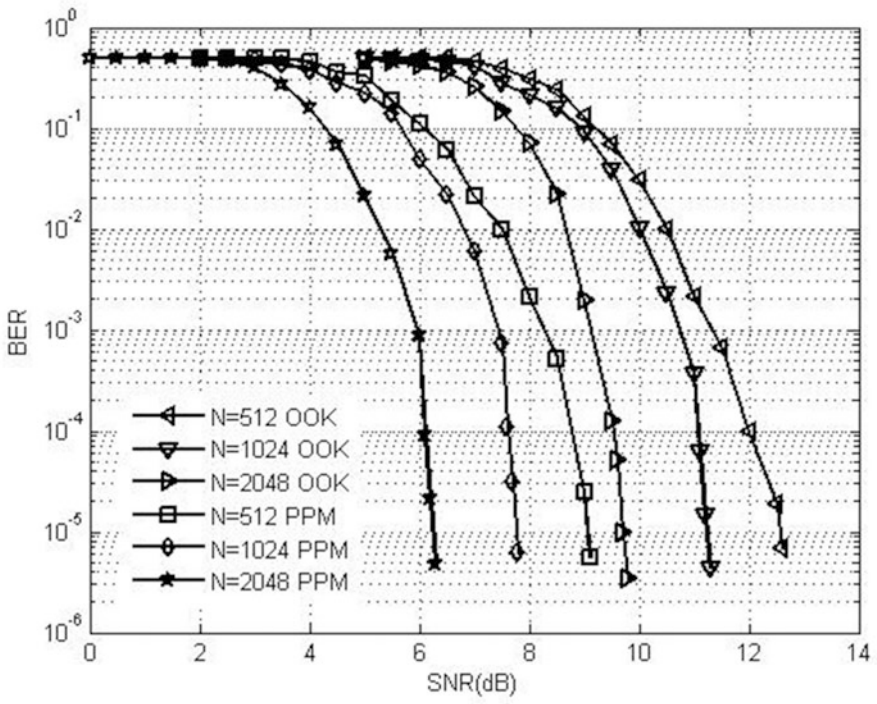


Fig. 18.18 PPM modulation with different code lengths [47]

rate characteristics by an order of magnitude and can effectively improve the bit-error rate performance of optical wireless systems [47].

Experimental Measurements

Experimental measurements were taken between September 2006 and January 2008. The experiment was conducted between two buildings in oblique transmission mode. The RS error-correction code adopted the RS (15, 9) code and the cyclic shortening RS (30, 10) code. The LDPC code rate is one-half and the code length is 128; the turbo code has a one-third bit rate and the interleaving length is 64.

Table 18.4 presents the system error rate before and after error correction on foggy, sunny, and cloudy days. Because of the low system error rate on sunny and cloudy days, no error rate was found after error correction. The bit-error rate measurement results of some rainy days are shown in Table 18.5, and the partial measurement results of a system in snowy weather are shown in Table 18.6.

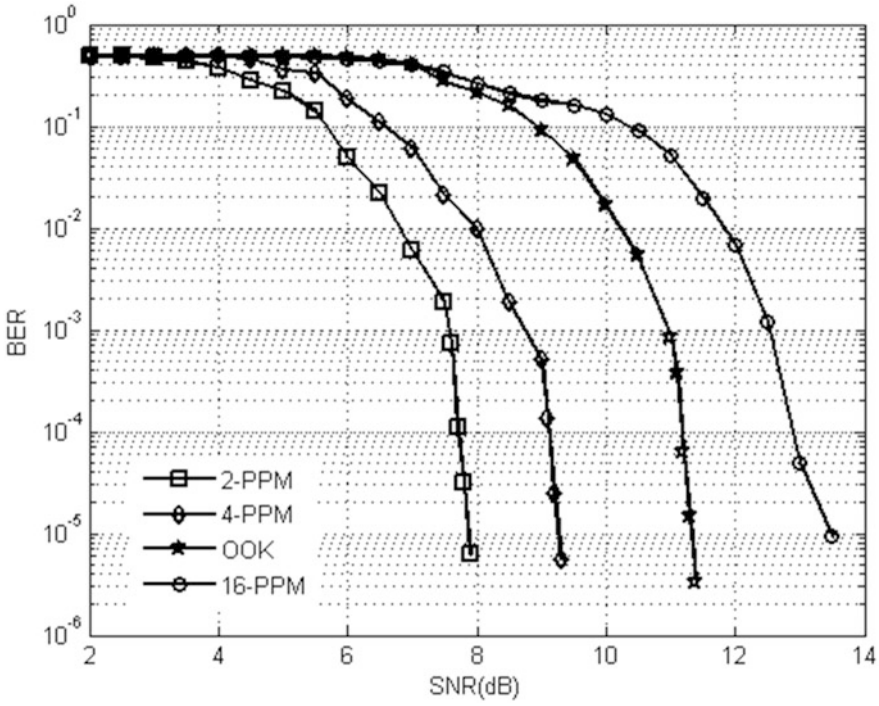


Fig. 18.19 Effect of the PPM timeslot number on the bit-error rate [47]

The experimental results show that, although the error-correction code can partially improve the performance of the optical wireless communication system, it is generally less effective in foggy and snowy weather, which leads to a significant increase in the error rate in snowy and foggy weather. In bad weather, the bit-error rate can reach 10^{-3} or even 10^{-2}. Most errors can be improved by using error-correction coding, e.g., 10^{-5}.

The error rate of turbo code and RS (30, 10) after error correction was below 10^{-5}. The bit-error rate was approximately 10^{-4} in light rain. After error correction using RS (15, 9) and the LDPC code, it was maintained at 10^{-6}. The average bit-error rate on sunny and cloudy days was 10^{-7}, but after error correction, the bit-error rate was better than 10^{-9} [50].

RS Codes

Figures 18.23 and 18.24 show the changes in the error rate before and after error correction using the two RS codes measured under different weather conditions. It can be seen from the figures that under light and moderate rain, the error rate of

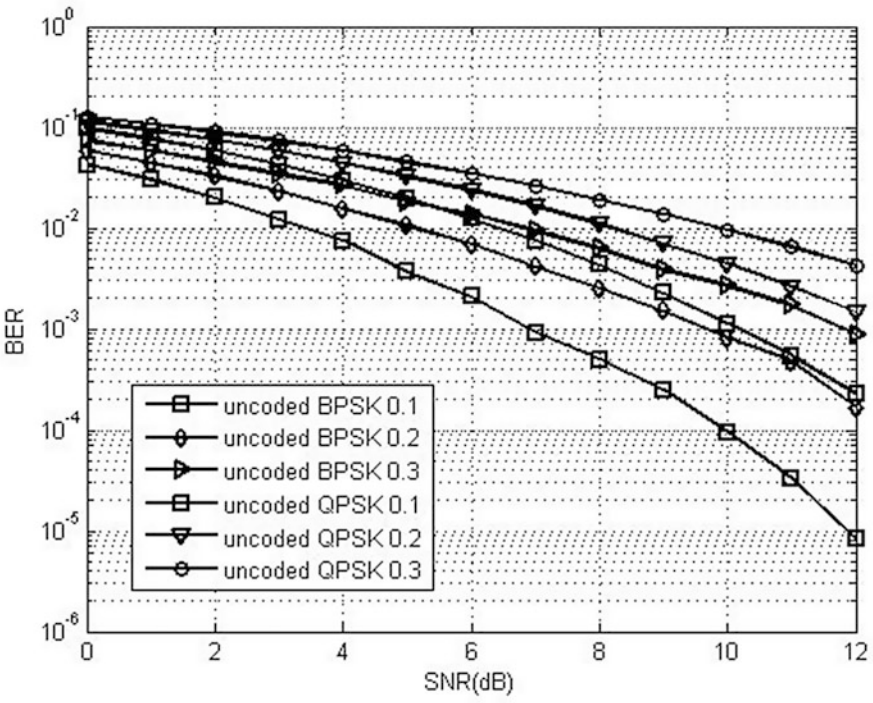


Fig. 18.20 Bit-error rate of an uncoded subcarrier system [42]

the RS (15, 9) code was significantly improved, from 10^{-4} to approximately 10^{-6}. However, for a high bit-error rate, the improvement was poor. In light to heavy rain, the error rate of the RS (30, 10) coding was reduced from 10^{-4} to 10^{-7}. Therefore, it can generally be used in rainy weather. In rough weather conditions, the bit-error rate of the system is generally between 10^{-3} and 10^{-2}. After error correction, the bit-error rate was effectively reduced and remained between 10^{-6} and 10^{-5} [50].

Turbo Codes

To further verify the improvement of the turbo code on the bit-error rate of optical wireless communication systems [51], we experimentally measured the bit-error rate in moderate rain, heavy rain, and rainstorm weather, and the results are shown in Fig. 18.25.

From Fig. 18.25, it can be seen that the turbo code improves the system, with a bit-error rate of about 10^{-4}; however, its correction ability is average for the system, with a bit-error rate of about 10^{-3}. Turbo codes are applicable to moderate rainfall and misty weather [50].

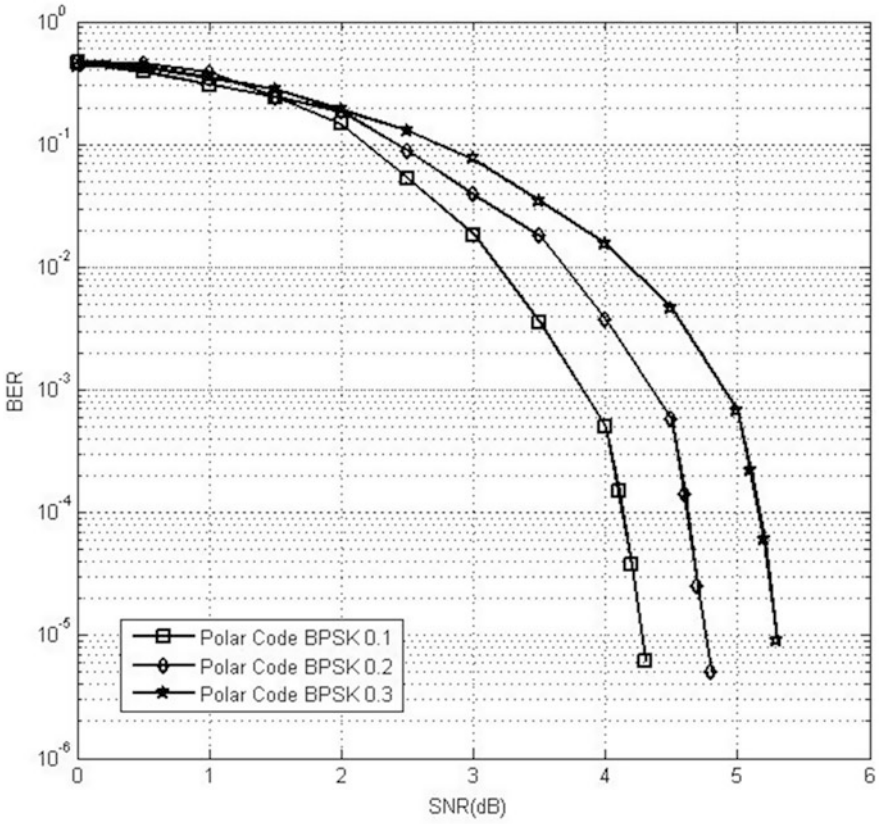


Fig. 18.21 Bit-error rate of the BPSK subcarrier system [47]

LDPC Codes

An LDPC code with a code length of 128 and code rate of 0.5 was adopted, and the measurement results are shown in Fig. 18.26 [50].

As shown in Fig. 18.26, when the error rate was between 10^{-4} and 10^{-3} under light and moderate rain conditions, the error rate after error correction was between 10^{-6} and 10^{-5}. In the case of a rainstorm, the error rate was between 10^{-3} and 10^{-2}, and the system error rate after error correction was still greater than 10^{-5}. LDPC coding is suitable for cases with a low bit-error rate. The system bit-error rate does not obviously improve in bad weather conditions [50].

Summary

This chapter systematically summarized the application of error-correction control coding in the field of optical wireless communication and introduced the work of Xi'an University of Technology in related fields [52], including the simulated and

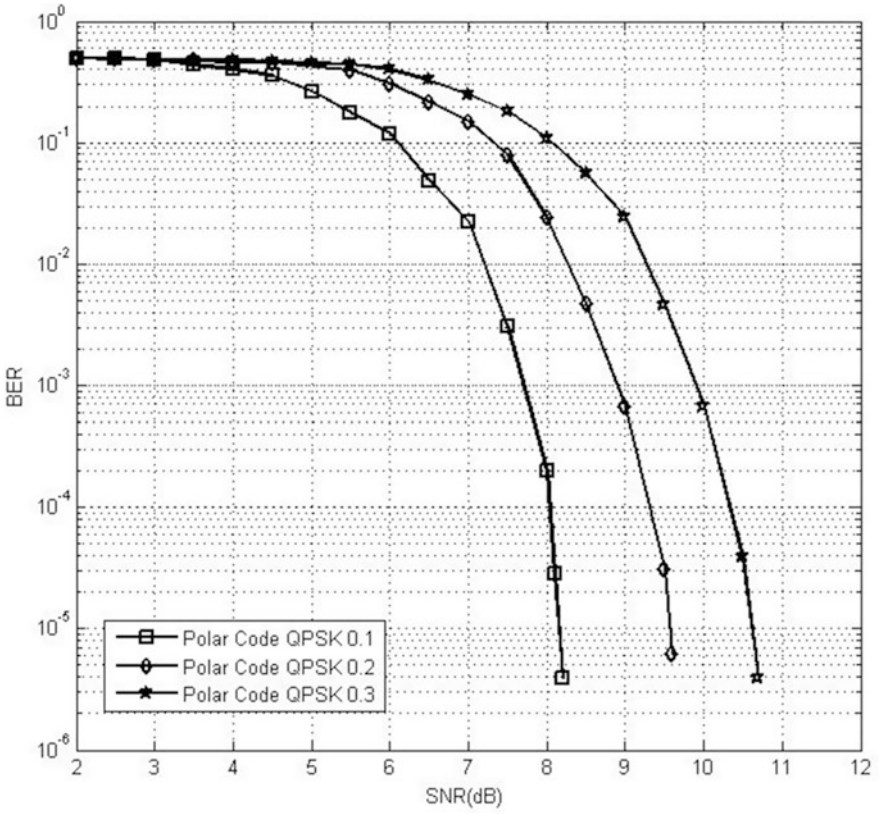


Fig. 18.22 Bit-error rate of a QPSK subcarrier system [47]

Table 18.3 Analysis of measured data from optical wireless subcarrier modulation systems [42]

	Sunny	Overcast	Rain
Symbols sent	10,240	12,288	13,824
Number of demodulation errors	122	12	147
Number of decoding errors	12	0	18
Demodulation error rate	0.0119	0.0000977	0.0106
Decoding error rate	0.00234	0	0.00261

experimental error-correction performance of RS codes, turbo codes, LDPC codes, and polarization codes under different modulation modes. These systems showed excellent performance in simulations and experimental measurements and effectively reduced the bit-error rate of optical wireless communication systems [51]. The development trends of error-correction coding in optical wireless communication systems can be summarized as follows [51, 52].

Table 18.4 Bit-error rate under foggy weather conditions [50]

Time	Weather	Visibility	Uncorrected BER	Error-correction method			
				LDPC	Turbo	RS (15, 9)	RS (30, 10)
2006-10-26	Dense fog	200 m	3.41×10^{-3}	5.74×10^{-5}		5.10×10^{-5}	
2006-10-27	Dense fog	200 m	3.35×10^{-3}	2.96×10^{-5}		2.72×10^{-5}	
2007-12-8	Rain and fog	150 m	7.3×10^{-3}		5.4×10^{-5}		3.2×10^{-5}
2007-12-9	Dense fog	200 m	2.6×10^{-3}	4.4×10^{-5}			7.6×10^{-6}
2007-12-10	Dense fog	400 m	1.3×10^{-3}			8.4×10^{-6}	2.5×10^{-6}
2006.9–2006-10	Sunny	15 km	$\leq10^{-7}$	$<10^{-9}$			
	Overcast	3 km	$\leq10^{-6}$	$<10^{-9}$			

Table 18.5 Bit-error rate under rainy weather conditions [50]

Time	Weather	Rainfall (mm)	Uncorrected BER	Error-correction method			
				LDPC	Turbo	RS (15, 9)	RS (30, 10)
2007-4-23	Light rain	0.2	2.2×10^{-4}	3.3×10^{-6}		3.0×10^{-6}	
2007-7-4	Heavy rain	21.8	6.8×10^{-4}		4.0×10^{-6}		2.4×10^{-6}
2007-6-17	Moderate rain	31.5	3.5×10^{-4}		8.4×10^{-7}		6.8×10^{-7}
2007-7-5	Rainstorm	34.1	2.4×10^{-3}	4.2×10^{-5}	8.8×10^{-6}	3.2×10^{-5}	
2007-8-9	Heavy rain	33.6	7.7×10^{-4}		2.2×10^{-6}		1.5×10^{-6}
2007-8-26	Heavy rain	76.9	8.2×10^{-3}			1.04×10^{-5}	
2007-8-30	Light to moderate rain	24.4	4.4×10^{-4}			5.0×10^{-6}	6.4×10^{-7}
2007-8-31	Light to moderate rain	2.4	3.8×10^{-4}		1.06×10^{-6}	4.8×10^{-6}	
2007-9-3	Light to moderate rain	10.1	4.0×10^{-4}			5.4×10^{-6}	8.6×10^{-7}
2007-9-7	Moderate to heavy rain	33.5	5.8×10^{-4}		3.4×10^{-6}		
2007-9-27	Light rain	7.0	2.6×10^{-4}	4.2×10^{-6}	8.3×10^{-7}		
2007-9-28	Light rain	8.7	3.2×10^{-4}			4.0×10^{-6}	
2007-9-29	Light rain	3.0	3.5×10^{-4}	4.3×10^{-6}		3.8×10^{-6}	
2007-10-8	Drizzle	0.2	7.6×10^{-5}			5.3×10^{-7}	
2007-10-9	Light rain	1.6	2.0×10^{-4}	3.2×10^{-6}	7.8×10^{-7}	3.0×10^{-6}	7.2×10^{-7}
2007-10-10	Light rain	6.6	2.6×10^{-4}	3.5×10^{-6}		3.1×10^{-6}	
2007-10-11	Light to moderate rain	12.0	4.8×10^{-4}	6.6×10^{-6}	1.8×10^{-6}		
2007-10-12	Light to moderate rain	28.6	4.0×10^{-4}			5.4×10^{-6}	8.8×10^{-7}
2007-10-27	Heavy rain	23.3	7.5×10^{-4}		2.4×10^{-6}		1.8×10^{-6}
2007-5-13	Moderate rain	13.6	5.0×10^{-4}	6.6×10^{-6}	7.5×10^{-7}	5.2×10^{-6}	
2007-10-15	Moderate rain	10.3	3.3×10^{-4}	5×10^{-6}		6×10^{-6}	
2007-7-2	Moderate rain	11.9	2.8×10^{-4}		5.4×10^{-7}		4.5×10^{-7}
2007-5-23	Light to moderate rain	9.2	4.7×10^{-4}			5.2×10^{-6}	6.1×10^{-7}
2007-5-24	Overcast and rainy	0.8	1.5×10^{-4}	1.8×10^{-6}		2.0×10^{-6}	
2007-7-4	Moderate to heavy rain	21.8	6.2×10^{-4}	8.4×10^{-6}		8.2×10^{-6}	
2007-5-23	Moderate rain	12.2	4.5×10^{-4}	5.4×10^{-6}	7.3×10^{-6}		

Table 18.6 Bit-error rate under snowy weather conditions [50]

Time	Weather	Snowfall (mm)	Uncorrected BER	Error-correction method			
				LDPC	Turbo	RS (15, 9)	RS (30, 10)
2007-1-3	Moderate snow	1.9	4.6×10^{-3}	7.8×10^{-5}		8.4×10^{-5}	
2008-1-11	Moderate snow	2.5	3.4×10^{-4}		2.1×10^{-6}		1.2×10^{-6}
2008-1-12	Major snow	5.6	6.2×10^{-3}	8.8×10^{-5}			3.6×10^{-5}
2008-1-18	Major snow	6.4	5.5×10^{-3}		3.8×10^{-5}		2.5×10^{-5}
2008-1-20	Major snow	6.0	4.4×10^{-3}		1.6×10^{-5}		9.2×10^{-6}
2008-1-21	Major snow	4.8	4.5×10^{-3}	6.3×10^{-5}		5.7×10^{-5}	

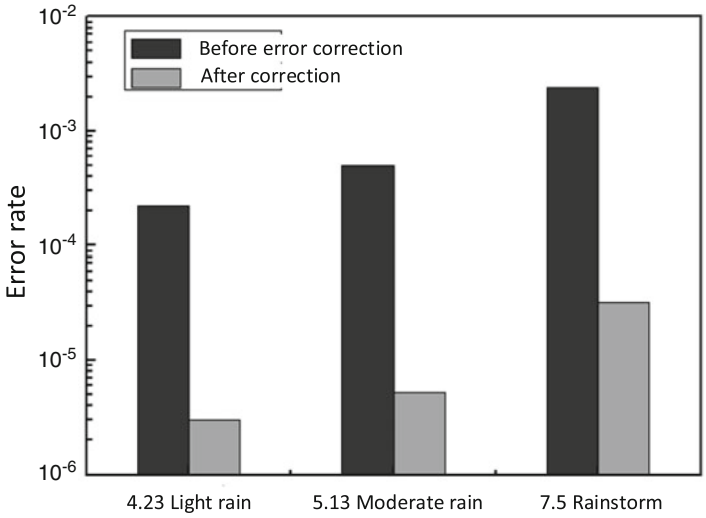


Fig. 18.23 RS (15, 9) code [50]

1. Adopt forward error correction to simplify system implementations. The effectiveness of a communication system is important; however, accurately transmitting information to the receiver is also an objective. Forward error correction can effectively improve the signal credibility and can correct error codes in the transmission process. Simplifying the system implementation is an important prerequisite for applying error-correction coding technology to optical wireless communication systems.
2. Explore codes that approach the Shannon limit and reduce the complexity of coding and decoding. To address the problems of coding efficiency and

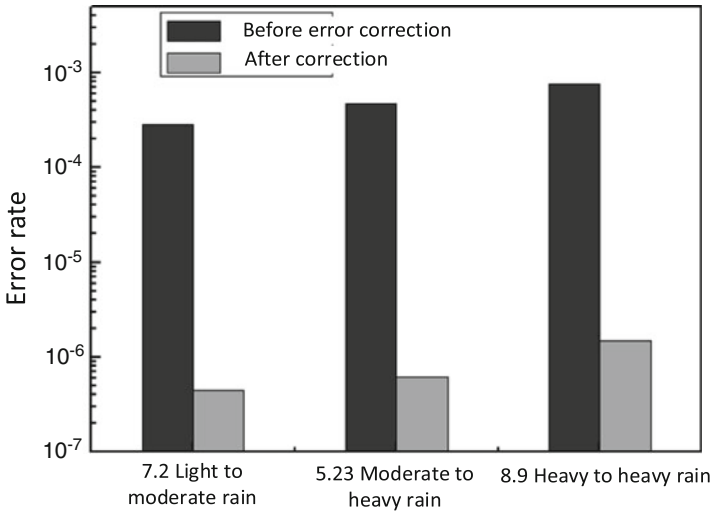


Fig. 18.24 RS (30, 10) code [50]

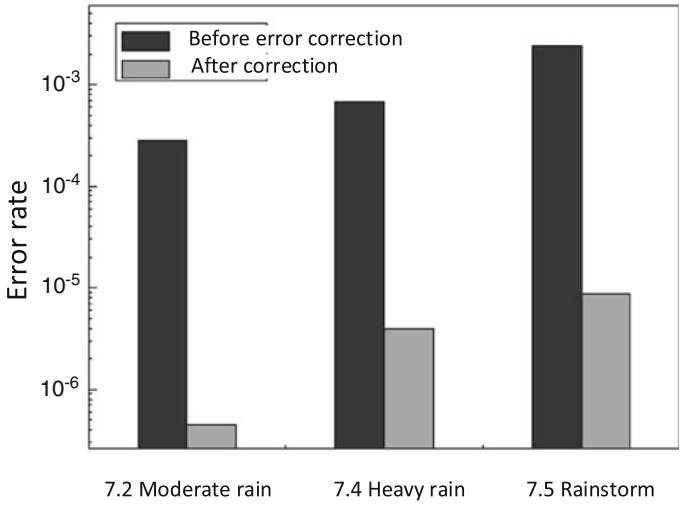


Fig. 18.25 Histogram of turbo code performance improvement [45]

computational complexity in current error-correction coding methods, a code with better performance was proposed. Polarization codes appeared late, and other error-correction codes have been applied to various fields. Therefore, problems have been encountered in practical applications. Polarization codes should be actively applied to practical communication systems. The coding and decoding algorithms of error-correction codes should be constantly optimized, their computational complexity reduced, and their implementation simplified.

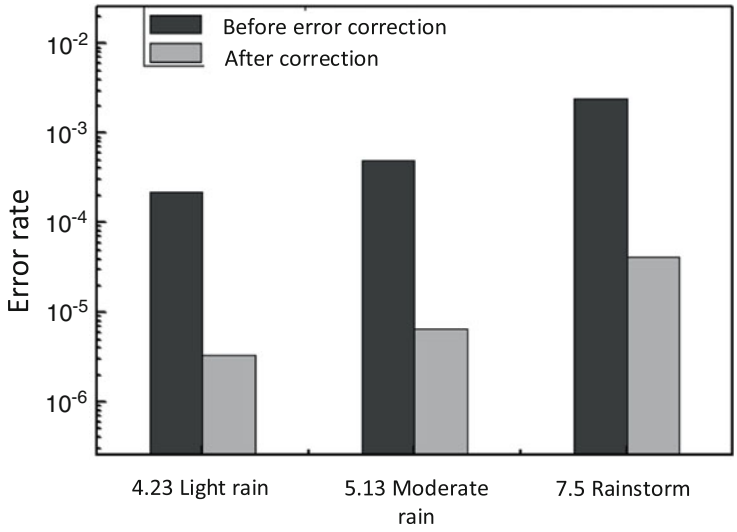


Fig. 18.26 Histogram of LDPC code performance improvement [44]

3. Make a reasonable compromise between coding complexity and error-correction performance and select the source modulation method that best matches the error-correction coding to improve efficiency.
4. Explore the effect of various error-correction codes on the performance of an actual experimental optical wireless system and make improvements.

Currently, research on the application of error-correction codes to optical wireless communication systems is mostly based on theoretical calculations and simulations. Although the performance of different error-correction codes applied to the system can be obtained, a combination of theory and practice can find the most suitable coding method for different communication systems, help reduce the deterioration of system performance encountered in the process of practical applications, and effectively improve the ability of optical wireless communication systems to suppress atmospheric turbulence.

The development of error-correction coding supports the suppression of the impact of turbulent channels on communication system performance and helps the promotion and application of optical wireless communication technology.

References

1. Ke X Z, Yin Z Y. Coding Theory in Wireless Laser Communication System [M]. Beijing: Science Press, 2009.
2. Ke X Z, Zhao L, Yin Z Y, et al. Experimental research on the error control of wireless laser communication [J]. Journal of Electronic Measurement and Instrument, 2009, 23(6): 18–27.
3. Shannon C E. A mathematical theory of communication [J]. The Bell system technical journal, 1948, 27(3): 379–427.

4. Hamming R W. Error detecting and error correcting codes [J]. In The Bell System Technical Journal, 1950, 29(2): 147–160.
5. Muller D E. Application of Boolean algebra to switching circuit design and to error detection [J]. Transactions of the IRE Professional Group on Electronic Computers, 1954, 3(3):6–12.
6. Reed I. A class of multiple-error-correcting codes and the decoding scheme [J]. Transactions of the IRE Professional Group on Information Theory, 1954, 4(4): 38–49.
7. Golay M J E. Notes on digital coding [J]. Proc. IEEE, 1949, 37(2): 657.
8. Elias P. Coding for noisy channels [J]. Proc.london Symp.information Theory London U.k, 1955, 3(1): 37–46.
9. Prange E. Cyclic Error-Correcting codes in two symbols [M]. Cambridge, Mass.: Air Force Cam-bridge Research Center, 1957.
10. Bose R C, Ray-Chaudhuri D K. On a class of error correcting binary group codes [J]. Information and control, 1960, 3(1): 68–79.
11. Hocquenghem A. Codes corecteurs d′erreurs [J]. Chiffres, 1959, 2: 147–156.
12. Reed I S, Solomon G. Polynomial Codes Over Certain Finite Fields [J]. Journal of the Society for Industrial & Applied Mathematics, 1960, 8(2):300–304.
13. Gallager R G. Low-Density Parity-check Codes [J]. IRE Trans., 1962, 8(1) :21–33.
14. Fang Y, Han G, Cai G, et al. Design guidelines of lowdensity paritycheck codes for magneticrecording systems [J]. IEEE Communications Surveys and Tutorials, 2018, 20(2): 1574–1606.
15. Berrou C, Glavieux A, Thitimajshima P. Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Code [C]//Proceedings of ICC'93-IEEE International Conference on Communications, Geneva, Switzerland, 23–26 May, 1993: 1064–1070.
16. Arikan E. Systematic Polar Coding [J]. IEEE Communications Letters,2011,15(8): 860–862.
17. Arikan. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels [J]. IEEE Trans. Inf. Theory, 2009, 55(7): 3051–3073.
18. Li C. Research on the performance of LDPC code and itsserialconca-tenatedcode [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2006: 71–73
19. Kang W M. Research of RS code technology in space optical communication [D]. Changchun: Jilin University, 2007.
20. Zhang W, Xu X Z, Zhang K, et al. Simulation analysis of RS concatenated code in ultra short wave communication and satellite communication channels [J]. Communications Technology, 2009, 42(2):27–29.
21. Yuan J G, Liang T Y, He L, et al. Research on the performance of the interleaved concatenated code for optical communication systems [J]. Semiconductor Optoelectronics, 2010, 31(2): 273–276.
22. Podemski R, Holubowicz W, Berrou C, et al. Hamming distance spectra of turbo-codes [J]. Annales des télécommunications, 1995, 50(9): 790–797.
23. Ohtsuki T, Sasase I, Mori S. Overlapping multi-pulse pulse position modulation in optical direct detection channel [C]//Proceedings of ICC'93-IEEE International Conference on Communications, Geneva, Switzerland, 23–26 May, 1993:1123–1131.
24. Zhu X, Kahn J M. Performance bounds for coded free-space optical communications through atmospheric turbulence channels [J]. IEEE Transactions on communications, 2003, 51(8): 1233–1239.
25. Anguita J, Djordjevic I, Neifeld M, et al. Shannon capacities and error-correction codes for optical atmospheric turbulent channels [J]. Journal of Optical Networking, 2005, 4(9): 586–601.
26. He X M, Li X F, Zhang D Y, et al. Performance for efficient error correction coding in Optical Wireless Communication [J]. Acta Photonica Sinica, 2008, 37(12): 2427–2429.
27. Wen L, Lei J, Wei B J. Design of concatenation of fountain and non-binary LDPC codes for satellite communications [C]//2010 2nd International Conference on Information Engineering and Computer Science, Wuhan, China, 25–26 Dec, 2010: 1–4.
28. Jiang X F, Zhao S H, Li Y J, et al. Design of interleaved R-S code of satellite-to-ground optical communication system [J]. Semiconductor Optoelectronics, 2011, 32(3): 401–404.

29. Habifj L, Dutton Z, Guha S. Polar coded optical communications with weak coherent states [C]// Lasers & Electro-optics, San Jose, CA, USA, 09–14, June, 2013: 1–2.
30. Wu Z F, Fischer J K, Lankl B. Experimental investigation of polar code performance for coherent UDWDM PONs [C]//2015 Optical Fiber Communications Conference and Exhibition (OFC), Los Angeles, CA, USA, 22–26 March, 2015: 1–3.
31. Fang, J F, Bi M H, Xiao S L, et al. Performance investigation of the polar coded FSO communication system over turbulence channel [J]. Applied optics, 2018, 57(25):7378–7384.
32. Iqbal S, Silva E, Yanko V, et al. An Experimental Demonstration of Rate- Adaptation Using Shaped Polar Codes for Flexible Optical Networks [J]. Journal of Lightwave Technology, 2019, 37(13):3357–3364.
33. Zhang H, Tu Q L, Cao Y, et al. Adaptive rate polar codes for atmospheric weakturbulencechannel [J]. Infrared and Laser Engineering, 2019, 48(7): 265–273.
34. Ao J, Tan X Y, Ma C B, et al. Design of free space optical communication system based on Raptor10 code [J]. Infrared and Laser Engineering, 2019, 48(9): 173–178.
35. Wen H, Cao Y, Peng X F, et al. MIMO polarization-coding method in free space optical communication [J]. Laser & Optoelectronics Progress, 2021, 58(19): 1906004.
36. Wen H, Cao Y, Peng X F, et al. Decoding method of relaxed polarization code in weak atmospheric turbulence channel [J]. Optical Communication Technology, 2021, 45(2): 36–40.
37. Lu X J, Shu B X, Lei J. Research on encoding and decoding technology of polar codes [J]. Radio Communications Technology, 2021, 47(6): 780–788.
38. Cao P, Rao Q, Yang J, et al. LDPC code with Dynamically adjusted LLR under FSO turbulence channel [J]. Journal of Physics Conference Series, 2021, 1920(1):1–8.
39. Dixit A, Jain V K. Analysis of LDPC codes in FSO communication system under fast fading channel conditions [J]. IEEE Open Journal of the Communications Society, 2021, 2(2): 1663–1673.
40. Du A Y, Ke X Z. Study and implementation of Reed-Solomon code in an atmospheric laser communication system [J]. Opto-Electronic Engineering, 2004, 31(S1): 44–47.
41. Du A Y. Research and implement of Reed-Solomon codes in atmosphere communication system [D]. Xi'an: Xi'an University of Technology, 2005.
42. Jia K J. Design and implementation of LDPC codes in atmosphere laser communication system [D]. Xi'an: Xi'an University of Technology, 2007.
43. Liu J. Research of turbo codes in atmosphere communication system [D]. Xi'an: Xi'an University of Technology, 2006.
44. Yin Z Y. Research on the characteristics of FSO generalized channel [D]. Xi'an: Xi'an University of Technology, 2009.
45. Sun Y S. Research and implementation of fountain code in low-voltage power line carrier meter reading system [D]. Xi'an: Xi'an University of Technology, 2009.
46. Shao J H, Ke X Z, Chen Q. A suitable polar coding modulation scheme for atmospheric weak turbulence channel [J]. Acta Electronica Sinica, 2016, 44(8): 1831–1836.
47. Chen Q. Research on polarization coding in Optical Wireless Communication system [D]. Xi'an: Xi'an University of Technology, 2014.
48. Niu K, Dai J S, Piao J N. Polar codes and polar processing for 6G wireless systems [J]. Journal on Communications, 2020, 41(5): 9–19.
49. Ke X Z, Lei S C, Shao J H, et al. Analysis on error rate of Optical Wireless Communication using subcarrier modulation based on polar code [J]. Infrared and Laser Engineering, 2015, 44(6): 1849–1853.
50. Wang X P. Experimental measure and analysis of the error code characteristic in the ALC system [D]. Xi'an: Xi'an University of Technology, 2008.
51. Ke X Z, Wu J L, Yang S J. Research progress and prospect of atmospheric turbulence for Optical Wireless Communication [J]. Chinese Journal of Radio Science, 2021, 36(3): 323–339.
52. Ke X Z, Yang S J, Wu J L, et al. Research progress of adaptive optics in Optical Wireless Communication system for Xi'an University of Technology [J]. High Power Laser and Particle Beams, 2021, 33(8): 30–52.

Optical Wireless MIMO Technology and Space–Time Coding

19

Contents

Introduction 616
Research Status of Optical MIMO Technology 616
Foreign Development Status 616
Domestic Development Status 618
Optical MIMO Systems 618
Composition of Optical MIMO Systems 619
Classification of Optical MIMO Technologies 619
Evolution from SISO to MIMO Channels 622
Urgent FSO-OFDM Problems 623
References 624

Abstract

The work of the previous authors is reproduced and summarized, including the authors' own views. Multiple-Input Multiple-Output (MIMO) technology makes full use of spatial resources and can suppress the influence of channel characteristics on optical-signal transmission. The evolution and development of optical wireless MIMO at home and abroad are described in detail from the perspective of improving the channel capacity and spectrum utilization of optical wireless communication. The principle of space–time coding is introduced. Finally, the key technical problems faced by optical wireless MIMO systems are summarized.

Keywords

Optical wireless multiple-input multiple-output system · Space–time coding · Research progress

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_19

Introduction

In recent years, interest has increased in using multiple-input multiple-output (MIMO) technology and its associated space–time coding techniques, through multicarrier modulation technology, smart antennas, and adaptive coding, to improve spectrum utilization [1–4]. MIMO utilizes spatial resources to increase the capacity and spectrum utilization of optical wireless systems [5–11].

Research Status of Optical MIMO Technology

Foreign Development Status

Optical wireless communication has been extensively studied since 1960 [5–9]. Institutions engaged in optical wireless communication research include the European Space Agency (ESA), United States Optical Communications Telescope Laboratory (CTL) [10–14], Japanese National Space Development Agency (NASDA) [15–19], and the United States National Aeronautics and Space Administration (NASA).

In 1980, high-bit-rate satellite communication experiments were conducted at the MIT Lincoln Laboratory [20–24]. In 1985, the United States Air Force conducted half-duplex optical-link communications between a ground station and the laser intersatellite transmission experiment (LITE) system [25–30]. Astro Terra developed a full-duplex optical wireless communication system, and the ESA verified the relevant unit technology of satellite optical communication [31–34]. In the 1990s, with the maturing of wavelength-division multiple access (WDMA), erbium-doped fiber amplifiers (EDFA), and other technologies, optical wireless communication improved greatly in terms of channel capacity and communication distance. In 1995, Japan conducted a satellite–ground optical communication experiment.

In 1984, Divsalar et al. [35] introduced coding and modulation technologies to the optical-communication field. Later, Reed–Solomon (RS) codes [36], trellis-coded modulation (TCM) codes [37], Bose–Chaudhuri–Hocquenghem (BCH) codes [38], turbo codes [39, 40], low-density parity-check (LDPC) codes [17], interleaving technology [41], and cascaded codes [42, 43] were applied to optical wireless communication.

In 1997, Wilson et al. [44] introduced the concept of multibeam transmissions to the field of optical wireless communication. The U.S. Jet Propulsion Laboratory (JPL) conducted several experiments on multibeam transmissions [45–48]. In October 1995, the effectiveness of multibeam transmissions in suppressing the atmospheric-turbulence effect was verified jointly by the United States and Japan [49]. Kim et al. [45] conducted an atmospheric-scintillation measurement experiment.

Theoretical models for beam transmissions in atmospheric turbulence have been presented in the literature [50, 51], but their practical application is still difficult. In 1996, Ibrahim [52] studied the performance of optical wireless communication using

receiver-diversity technology. In 2000, Bell Laboratories developed its prototype system [53]. In 2002, the JPL conducted a 46.8-km multibeam optical transmission experiment [54] and, using spatial-diversity reception, conducted an outfield test [55–57].

In 2001, the JPL and the laser-communication research group of Stanford University conducted an in-depth study on diversity reception and optical-signal detection and achieved good results [58]. Haas et al. [59] proposed a space–time coding criterion for atmospheric laser communication, based on a log-normal attenuation atmospheric-channel (weak undulating turbulence) model and deduced its bit-error rate (BER) performance limit.

Coherent detection requires a very stable frequency for the receiver's local oscillation light and the transmitted light, which affects its practical application [60, 61]. Reference [62] provides the space–time coding criterion and BER performance limit, which are not suitable for optical communication systems with direct detection.

In 2002, Zhu et al. [63, 64] applied a diversity-reception technique in radio-frequency (RF) communication to optical wireless communication. In 2003, Simon et al. [65] modified and improved the Alamouti code. Alqudah et al. [66] proposed orthogonal space–time coding for indoor optical wireless channels. In 2004, Pan et al. [67] found that multibeam transmissions could effectively suppress the light-intensity fluctuations caused by atmospheric turbulence.

Wilson et al. [68] studied the BER performance of a quantized pulse-position modulation system (QPPM) with modulated MIMO optical communication. Brandt-Pearce et al. [69] analyzed the BER performance limit using PPM modulation under the same conditions. In 2005, Simon et al. [70] proposed a space–time coding scheme based on direct detection (intensity-modulation direct detection, IM/DD). In 2006, Angulta et al. [71] found that multibeam modes could achieve a higher signal-to-noise gain than single-beam transmissions. In 2007, Angulta et al. [72] analyzed the effects of the transmission distance, turbulence conditions, transmitter spacing, and receiver aperture on the performance of a communication system.

Navidpour et al. [73] derived a formula for calculating the BER when the transceiver simultaneously used diversity technology. Cao et al. [74] designed a PPM optical MIMO system. García-Zambrana analyzed the BER performance of Simon's space–time block code in strongly turbulent channels [75], and found that the space–time block code could suppress atmospheric turbulence at the cost of reduced multiplexing gain.

In 2008, Cvijetic et al. [76] derived performance bounds for optical communication links with multibeam transmit–multidetection receive. Reference [77] analyzed the interruption capacity of MIMO systems with IM/DD in Poisson fading channels, but did not consider their optical wireless communication performance with thermal-noise IM/DD modulation [74]. References [77, 78] investigated the performance of the system in different cases, but did not investigate how to achieve the maximum channel capacity for MIMO transmissions using the optimal transmission mechanism.

Safari et al. [79] studied the performance of on–off keying (OOK) modulated repetition codes and space–time codes. The BER performance of the two was very close in strong turbulence; in weak turbulence, the former outperformed the latter. Zaidi et al. [80] proposed a cross-coding scheme that combined a medium access control (MAC) layer and orthogonal space–time coding.

In 2009, Ntogari et al. [81] studied the performance of space–time coding in an indoor environment and found that space–time block codes improved the channel capacity. In 2010, Bayaki et al. [82] used two lasers and arbitrary-detector reception to obtain a closed expression of the pairwise error probability of space–time coding. García-Zambrana et al. [83] proposed a space–time trellis code and analyzed the BER performance in a strong turbulent channel.

Domestic Development Status

In 1999, the Northwest Institute of Nuclear Technology found that the light intensity of a multibeam transmission was twice as high as that of a single-beam transmission [84]. In 2003, the University of Electronic Science and Technology of China studied and verified the characteristics of multibeam transmissions to suppress the atmospheric-turbulence effect [85, 86]. In 2004, the National University of Defense Technology [87–89] studied the channel model and space–time coding of multibeam transmissions, and the Shanghai Institute of Optics and Fine Mechanics implemented a communication system based on multiple transmitters [90]. The system adopted transmitting diversity instead of receiving diversity and space–time coding.

In 2005, the Huazhong Institute of Optoelectronic Technology [91] and Yanshan University [92] studied diversity technology. Similar studies were reported by Anhui University [93] and Beihang University [94]. In 2006, Wuhan University verified the performance of optical wireless space–time coding through simulation [95]. Since 2007, the team of Professor Ke Xizheng from Xi'an University of Technology has investigated space–time coding, and in 2019, they conducted a 100-kilometer multibeam-transmission experiment. Subsequently, Huazhong University of Science and Technology [96, 97] and Changchun University of Science and Technology [98, 99] also researched space–time coding.

Optical MIMO Systems

Owing to the differences between radio-frequency communication and wireless laser communication, the theory and model of the original MIMO technology significantly differ from those of wireless laser-communication systems [14]. Optical MIMO can effectively suppress the atmospheric-turbulence effect and improve the channel capacity.

Composition of Optical MIMO Systems

An optical MIMO system is an organic fusion of coding, modulation, and diversity, as shown in Fig. 19.1.

Optical MIMO divides the source data into groups according to certain rules, and obtains multiple modulation symbols after constellation mapping. Then, the corresponding coding matrix is obtained using space–time coding. At the receiving end, the optical signal is converted into the corresponding electrical signal. After decoding, it is converted into a multichannel sub-stream to improve the channel capacity and antifading ability of the system [100].

Classification of Optical MIMO Technologies

Diversity Technology

Diversity is a basic technology for suppressing fading in wireless communication systems [14]. It suppresses channel fading and improves the system reliability [101]. Diversity technology receives multiple signal sequences from many different channels and its fading usually shows some independence. Spatial diversity [14] is realized using multiple antennas, and its performance is better than that of time and frequency diversity.

The anti-jamming ability of receiving diversity is limited, and must be supplemented by other anti-jamming methods [102]. Time diversity [14] combines time interweaving with channel coding; it has a limited effect on slow fading, but is very effective for fast fading. Time diversity contains a large number of redundant signals caused by error-correcting codes in the signals at the receiving end [102]. In spatial diversity [14], multiple antennas are staggered at a certain distance in space such that the signals received by each antenna are not correlated [103].

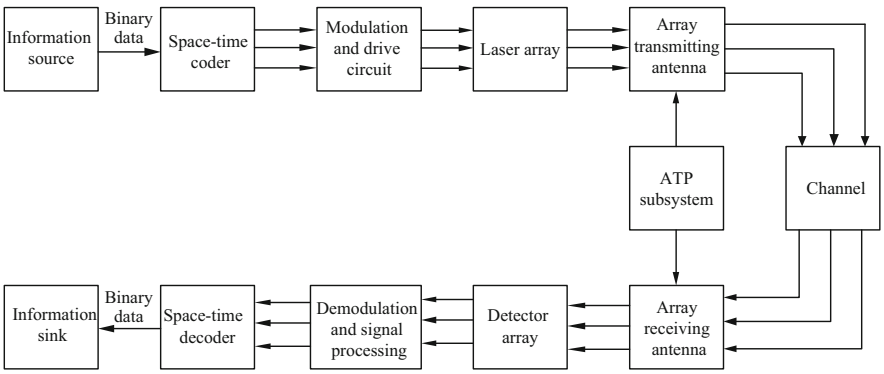


Fig. 19.1 Composition of an optical MIMO communication system

Diversity Combining Technology

Introducing diversity technology into optical wireless communication can effectively suppress the fading caused by atmospheric turbulence [104]. Diversity technology can multiplex the same signal to prevent all signals from being simultaneously affected by atmospheric fading [105].

Equal-Gain Combining-Receiving Diversity

As shown in Fig. 19.2, assuming that the number of receiving antennas is $N \geq 2$, the total receiving aperture area of the N receiving antennas is SR, the aperture area of each receiving antenna is SR/N, and it is an equal-gain combination. It is assumed that the receiving aperture area is equal to that of a nondiversity single-input single-output (SISO) system.

Equal-Gain Combining of Transmit and Receive Diversity

As shown in Fig. 19.3, we assume that the number of transmitting antennas $M \geq 2$, number of receiving antennas $N \geq 2$, area of each transmitting aperture is AT/M, transmitting power is $1/M$ of the transmitting power of a single-ended receiving system without diversity, and the area of each receiving antenna is AR/N.

Reference [106] proves that the single-transmit receive diversity in Fig. 19.2 is equivalent in effect to the spatial diversity in Fig. 19.3, which uses $M \times N$ receiving antennas, and the power gain is the same.

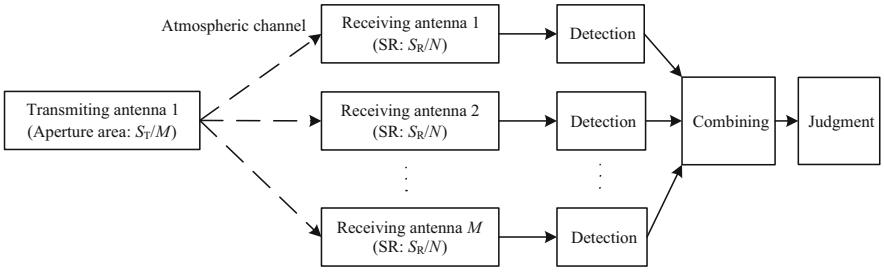


Fig. 19.2 Equal-gain combining-receiving diversity block diagram [105]

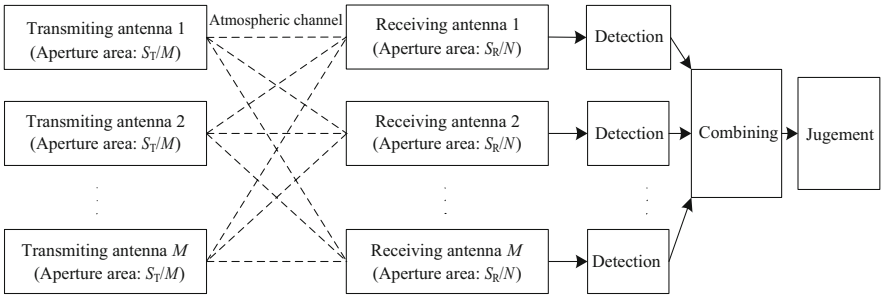


Fig. 19.3 Equal-gain combining transmit and receive diversity block diagram [105]

Spatial Multiplexing Technology in MIMO Systems

A schematic diagram of spatial multiplexing is shown in Fig. 19.4. Various layered space–time codes use spatial multiplexing to improve the channel capacity. The main receiver-side decoding algorithms for spatial multiplexing include the zero-forcing algorithm (ZF), minimum mean-squared-error algorithm (MMSE), and maximum-likelihood algorithm (ML) [107].

Spatial multiplexing means each sub-stream is sent from a different antenna, which is similar to diversity transmission, except that the full ergodic capacity can be obtained on the independent and identically distributed Rayleigh channels [108]. However, it cannot provide the same diversity gain as MIMO diversity and requires a reasonable compromise between rate and reliability. The essence of MIMO systems is the compromise between the system error probability and the system data-transmission rate [14]. Increasing the spatial-diversity gain results in a lower spatial multiplexing gain.

The Bell Laboratories layered space–time (BLAST) algorithm [14] is divided into the diagonal BLAST (D-BLAST) and vertical BLAST (V-BLAST) algorithms. V-BLAST encodes, maps, and interleaves M bit streams, and then transmits them independently to M antennas, which maximize diversity; all information streams are detected independently [109–112].

D-BLAST is basically the same as V-BLAST in terms of coding and modulation methods and sequences, except that the signal is divided into different signal blocks after processing. These signal blocks are assigned to different antennas for simultaneous transmission; thus, the effect of the entire signal transmission is not significantly affected when the transmission is not too ideal on one line. However, the system complexity correspondingly increases [113, 114]. BLAST can make full use of scarce spectrum resources; however, it is less feasible for downlinks because the number of transmitting antennas must be less than or equal to the number of receiving antennas. Moreover, because all antennas transmit different signals, the information in each link is unique, and when a link is broken, the signal integrity is not guaranteed [115, 116].

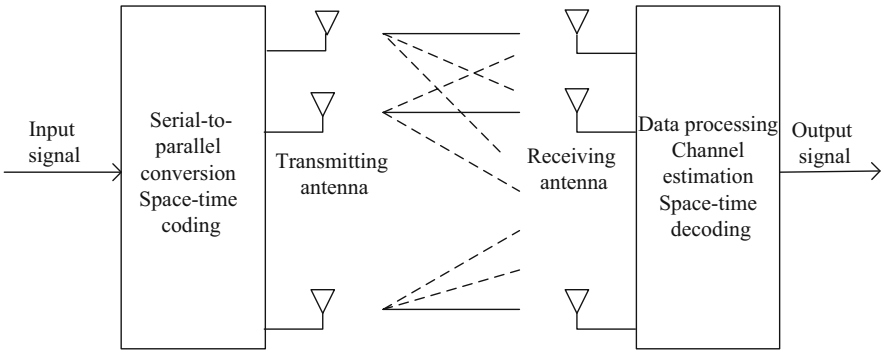


Fig. 19.4 Spatial multiplexing diagram

Evolution from SISO to MIMO Channels

The SISO wireless channel model is the basis for all derivative channels and rarely involves the study of spatial characteristics. The key point that distinguishes the MIMO channel model from SISO is the spatial characteristics of the wireless channel [117]. The vector model of the MIMO system is derived according to the spatial characteristics, and based on the SISO-channel vector model [118].

The SISO wireless channel system is a single transmitting and receiving antenna-communication system; its channel impulse response is

$$h(t,\tau) = \sum_{i=1}^{L} \alpha(t,\tau) e^{j\varphi_i(t,\tau)} \delta[\tau - \tau_i(t)], \tag{19.1}$$

where $\varphi(t, \tau)$ denotes the sum of the phase shifts of the multipath propagation channels.

The channel model of a system with one antenna at the transmitter and M receiving antennas at the receiver is called SIMO. According to reference [1], a SIMO channel can be combined as M SISO scalar channels, and its channel impulse response is

$$\boldsymbol{h}(t,\tau) = h_1(t,\tau), h_2(t,\tau), \ldots, h_M(t,\tau)^{\mathbf{T}}, \tag{19.2}$$

where $h_i(t, \tau)$ is the impulse response of the ith SISO channel. Then, the impulse response of the SIMO channel is

$$\boldsymbol{h}(t,\tau) = \sum_{i=0}^{l-1} \alpha(\theta_i,\varphi_i) \beta_i^{j(2\pi f_i \tau_i + \varphi_i)} \delta[\tau - \tau_i(t)]. \tag{19.3}$$

Here, β_i denotes the path gain of the ith subchannel, τ_i denotes the path delay of the ith subchannel, $\alpha(\theta_i, \varphi_i)$ denotes the direction of arrival of the ith subchannel, f_i denotes the Doppler frequency shift, φ_i denotes the additional phase shift, and $\alpha(\theta, \varphi)$ denotes the array steering vector. The expressions are as follows:

$$\boldsymbol{a}(\theta,\varphi) = 1, \alpha_1(\theta,\varphi), \ldots, \alpha_{M-1}(\theta,\varphi)^{\mathrm{T}}, \tag{19.4}$$

where the ith component is

$$\alpha(\theta_i,\varphi_i) = e^{j2\pi(x_i \sin\theta\cos\varphi + y_i \sin\theta\sin\varphi + z_i\cos\theta)}. \tag{19.5}$$

Comparing the SIMO and SISO channels, the scalar channel expands to the vector channel, thus showing the spatial channel characteristics of the system.

The MISO channel employs N transmit antennas and a single receive antenna. Considering the channel as a vector channel formed by the combination of N SISO scalar channels [120], the MISO channel impulse response can be obtained as

$$\boldsymbol{h}(t,\tau) = h_1(t,\tau), h_2(t,\tau), \ldots, h_N(t,\tau)^{\mathrm{T}}, \tag{19.6}$$

where $h_i(t, \tau)$ is the impulse response of the ith subchannel. The impulse response of the MISO channel is

$$\boldsymbol{h}(t,\tau) = \sum_{i=0}^{l-1} \alpha(\theta_i, \varphi_i)\beta_i^{j(2\pi f_i \tau_i + \varphi_i)} \delta\left[\tau - \tau_i(t)\right], \tag{19.7}$$

where $\alpha(\theta_i, \varphi_i)$ denotes the wave direction.

According to the formation ideas of SIMO and MISO, a MIMO channel with N and M respective transmit and receive antennas can be regarded as a matrix channel composed of $N \times M$ SISO scalar channels.

$$\boldsymbol{h}(t,\tau) = \begin{pmatrix} h_{11}(t,\tau) & h_{12}(t,\tau) & \cdots & h_{1n}(t,\tau) \\ h_{21}(t,\tau) & h_{22}(t,\tau) & \cdots & h_{2n}(t,\tau) \\ \vdots & \vdots & \ddots & \vdots \\ h_{m1}(t,\tau) & h_{m2}(t,\tau) & \cdots & h_{mn}(t,\tau) \end{pmatrix}, \tag{19.8}$$

where, $\boldsymbol{h}_{ij}(t, \tau)$ is the impulse response of the subchannel (SISO scalar channel) between the ith transmitting antenna and the jth receiving antenna.

Urgent FSO-OFDM Problems

Owing to the difference between radio-frequency orthogonal frequency-division multiplexing (RF-OFDM) and free-space optical orthogonal frequency-division multiplexing (FSO-OFDM), as well as the shortcomings of FSO technology, many problems can be solved by combining them.

1. RF-OFDM and optical OFDM have the same principle [119]; however, there are essential differences, and the RF-OFDM system model cannot simply be applied to optical OFDM.
2. Source and channel coding [120–123] can significantly improve the antiinterference capability of communication systems. Because FSO cannot directly transmit signals in a complex manner [124], the coding method needs to be improved.
3. The FSO-OFDM system has carrier, symbol, and sample synchronization. Because the OFDM signal is composed of multiple mutually orthogonal subcarrier signals, ensuring this orthogonality is crucial for OFDM systems [125 –128].
4. Because the FSO-OFDM channel is affected by the scattering of particles, such as dust, raindrops, and fog, a cyclic-protection interval is introduced in OFDM. Simple frequency-domain equalization can be used at the receiving end to eliminate multipath interference caused by atmospheric scattering. However, the exact channel frequency response on each subcarrier must be known before frequency-domain equalization.

OFDM technology has just appeared in the field of optical wireless research, and its theory and technology are imperfect. In-depth research is urgently needed to better exploit the potential of their respective technologies.

References

1. Liberti J C, Rappaport T S. Smart antennas for wireless communications: IS-95 and third generation CDMA applications [M]. Prentice Hall PTR, 1999.
2. Huang X H, Wu S L. DOA and Doppler frequency estimation algorithm based on uniform circular array [J]. Electronics Journal. 2001, 29(5):619–624.
3. Khalighi M A, Brossier J M, Jourdain G, et al. Water Filling Capacity of Rayleigh MIMO Channels[C]. 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. 2001, A: 155–158.
4. Gao J. Application of MIMO technology in UWB system [D]. Xi'an: Xi'an Electronic Technology University, 2008.
5. Djordjevic I B, Vasic B, Neifeld M A. LDPC Coded OFDM Over the Atmospheric Turbulence Channel [J]. Optics Express, 2007, 15(10):6336–6350.
6. Paterson C. Atmospheric Turbulence and Orbital Angular Momentum of Single Photons for Optical Communication [J]. Physical Review Letters, 2005, 94(15):153901-1-4.
7. Polishuk A, Arnon S. Optimization of A Laser Satellite Communication System with An Optical Preamplifier [J]. Journal of the Optical Society of America A, 2004, 21(7):1307–1315.
8. Ricklin J C, Davidson F M. Atmospheric Optical Communication with A Gaussian Schell Beam[J]. Journal of the Optical Society of America A, 2003, 20(5):856–866.
9. Tang Y, Shieh W, Yi X W, et al. Optimum Design for RF-to-Optical Up-Converter in Coherent Optical OFDM Systems [J]. IEEE Photonics Technology Letters, 2007, 19(7):483–485.
10. Wang H Q, Ke X Z. Free space optical communication based on vertical layered space-time coding [J]. China Laser, 2008, 35(6):874–878.
11. Chen J, Huang D X, Yuan X H. Characteristics analysis of free space optical communication system based on dual-mode Turbo code [J]. China Laser, 2006, 33(11):1532–1536.
12. Ke X Z, Ying Z Y. Coding theory in wireless laser communication system [M]. Beijing: Science Press, 2009.
13. Ke X Z, Deng L J. Theory of partially coherent optical transmission in Optical Wireless Communication [M], Beijing: Science Press, 2018.
14. Ke X Z. Space-time coding theory in wireless optical MIMO systems [M]. Beijing: Science Press, 2014.
15. Ke X Z, Dong K. Optical Wireless Communication [M]. Springer Press, 2022
16. Sahuguede S, Julien-Vergonjanne A, Cances J P. Soft Decision LDPC Decoding over Chi-square Based Optical Channels [J]. Journal of Lightwave Technology, 2009, 27(16): 3540–3545.
17. Sahuguede S, Fafchamps D, Julien-Vergonjanne A, et al. LDPC Code Design and Performance Analysis on OOK Chi-Square-Based Optical Channels [J]. IEEE Photonics Technology Letters, 2009, 21(17):1190–1192.
18. He X M, Li X F, Zhang D Y, et al. Performance analysis of efficient error correction coding technology in Optical Wireless Communication [J]. Photonics Journal, 2008, 37(12): 2427–2429.
19. Xiang P, Wang R. Research on Dynamic Wavelength Routing and Slot Allocation Algorithms in WDM-TDM Optical Networks [J]. Journal of Electronics and Informatics, 2009, 31(3): 679–683.
20. Zhou G T, Zhang X G, Xi L X, et al. Experimental study on two-stage polarization mode dispersion adaptive compensation in 40Gb/s optical time division multiplexing system [J]. Journal of Optics, 2005, 25 (1):15–23.

21. Pan W Q, Liu L R, Liu X M. Imaging optical communication system and codec scheme [J]. China Laser, 2006, 33(2):213–223.
22. Djordjevic I B. LDPC-coded MIMO Optical Communication Over the Atmospheric Turbulence Channel Using Q-ary Pulse-position Modulation [J]. Optics Express, 2007, 15 (16): 10026–10032.
23. Wilson S G,Brandt-Pearce M,Cao Q L, et al. Free-Space Optical MIMO Transmission with Q-ary PPM [J]. IEEE Transactions Communications, 2005, 53(8): 1402–1412.
24. Awan M S, Brandl P, Leitgeb E, et al.Transmission of High Data Rate Optical Signals in Fog and Snow Conditions [C]. International Conference on Wireless Communication, 2009: 702–706.
25. Feng B, Fan G, Min Y. Dynamic probabilistic packet marking based on PPM [C]. Proceedings of the 2009 2nd Pacific-Asia Conference on Web Mining and Web-Based Application, 2009: 289–292.
26. Javornik T, Jelovčan I, Muhammad S S, et al. Simplified Soft Value Extraction for M-PPM-modulated Signals in FSO systems [J]. International Journal of Electronics and Communications, 2009, 63: 595–599.
27. Huang J F, Yang C C, Huang C M. On Analyzing Quasi-Cyclic LDPC Codes Over Modified Welch–Costas-Coded Optical CDMA System [J]. Journal of Lightwave Technology, 2009, 27(12): 2150–2158.
28. Kiasaleh K. Turbo-Coded Optical PPM Communication Systems [J]. Journal of Lightwave Technology, 1998, 16(1):18–30.
29. Kim J Y, Poor H V. Turbo-Coded Optical Direct-Detection CDMA System with PPM Modulation [J]. Journal of Lightwave Technology, 2001, 19(3):312–327.
30. Liu D P. Research on high precision alignment information acquisition method in space optical communication [D]. Chengdu: University of Electronic Science and Technology, 2005.
31. Laurent B, Planche G. SILEX Overview after Flight Terminals Campaigh [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1997, 2990: 10–26.
32. Fenner W R. Future Trends in Crosslink communications [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1993, 1866:1–8.
33. Oppenhauser G. SILEX Program Status-a Major Milestone is Reached [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1997, 2990:2–9.
34. Bondurant R S, Boroson D M. Overview of Lasercom Program at Lincoln Laboratory [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1995, 2381:2–3.
35. Divsalar D, Gagliardi R M, Yuen J H. PPM Performance for Rees-solomon Decoding Over an Optical- RF Relay Link [J]. IEEE Transactions on Communications, 1984, 32(3):302–305.
36. Koh Y T, Davidson F. Interleaved Concatenated Coding for the Turbulent Atmospheric Direct Detection Optical Communication Channel [J]. IEEE Transactions on Communications, 1989, 37(6):648–651.
37. Park H, Barry J R. Trellis-Coded Multiple-Pulse-Position Modulation for Wireless Infrared Communications [J]. IEEE Transactions on Communications, 2004, 52(4): 643–651.
38. Seki K, Mikami K, Katayama A, et al. Single-chip FEC Codec Using a Concatenated BCH Code for 10 Gb/s Long-Haul Optical Transmission Systems [C]. IEEE Custom Integrated Circuits Conference, 2003: 279–282.
39. Cai Y, Morris J M, Adali T, et al. On Turbo Code Decoder Performance in Optical-Fiber Communication Systems With Dominating ASE Noise [J]. Journal of Lightwave Technology, 2003, 21(3):727–734.
40. Chitra K, Ravichandran V C. Turbo-coding of Coherence Multiplexed Optical PPM CDMA System with Balanced Detection [J]. American Journal of Applied Sciences, 2007, 4(5): 264–268.
41. Davidson F M, Koh Y T. Interleaved Convolution Coding for the Turbulent Atmospheric Optical Communication Channel [J]. IEEE Transactions on Communications, 1988, 36(9): 993–1003.

42. Magarini M, Essiambre R J, Basch B E, et al. Concatenated Coded Modulation for Optical Communications Systems [J]. IEEE photonics technology letters, 2010, 22(16): 1244–1246.
43. Djordjevic I B, Xu L, Wang T. Reverse Concatenated Coded Modulation for High-Speed Optical Communication [J]. IEEE Photonics Journal, 2010, 2(6): 1034–1039.
44. Wilson K E, Leatherman P R, Cleis R, et al. Results of the Compensated Earth-Moon-Earth Retro-reflector Laser Link (CEMERLL) experiment [R]. TDA Progress Report 42-131, 1997:1–13.
45. Kim I. I, Hakakha H, Adhikari P, et al. Scintillation Reduction Using Multiple Transmitters [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1997, 2990: 102–113.
46. Kim I I, Woodbridge E, Chan V, et al. Scintillation Measurements Performed During the Limited-visibility Lasercom Experiment [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1998, 3266: 209–223.
47. Kim I I, Mitchell M, Korevaar E. Measurement of Scintillation for Free Space Laser Communication at 785nm and 155 nm [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1999, 3266:49–62.
48. Kim I I, Koontz J, Hakakha H, et al. Measurement of Scintillation and Link Margin for the Terralink Laser Communication System [C]. Proceedings of SPIE - The International Society for Optical Engineering, 1998, 3232:100–118.
49. Wilson K E, Jeganathan M, Lesh J R, et al. Results from Phase-l and Phase-2 GOLD Experiments [R]. TDA Progress Report 42-133. 1997:1–13.
50. Strohbehn J W. Laser Beam Propagation in the Atmosphere [M]. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978.
51. Tatarskii V I. Wave Propagation in a Turbulent Medium [M]. New York: McGraw, 1961.
52. Ibrahim M M, Ibrahim A M. Performance Analysis of Optical Receivers with Space Diversity Reception [J]. IEEE Proceedings - Communications, 1996, 143(6):369–372.
53. Nykolak G, Szajowski P F, Cashion A, et al. A 40Gbps DWDM Free Space Optical Transmission Link Over 4.4 km [C]. Proceedings of SPIE - The International Society for Optical Engineering, 2000, 3932:16–23.
54. Biswas A, Wright M W. Mountain-Top-to-Mountain-Top Optical Link Demonstration [R]. Part I. IPN Progress Report, 2002: 42–149.
55. Vilnrotter V, Lau C W, Andrews K, et al. Real-time Combining of Optical Array Signals [C]. Proceedings of SPIE - The International Society for Optical Engineering, 2006, 6105:50–60.
56. Vilnrotter V A, Lau C W, Andrews K, et al. Two-element Optical Array Receiver Concept Demonstration [C]. Proceedings of SPIE - The International Society for Optical Engineering, 2005, 5712:225–239.
57. Vilnrotter V, Lau C W, Srinivasan M, et al. Optical Array Receiver for Communication Through Atmospheric Turbulence [J]. Journal of Lightwave Technology, 2005, 23(4): 1664–1675.
58. Vilnrotter V, Srinivasan M. Optical Communications through Atmospheric Turbulence Using Photodetector Arrays [C]. Proceedings of SPIE - The International Society for Optical Engineering, 2001, 4272:282–292.
59. Haas S M, Shapiro J H. Capacity of Optical Wireless Communications [J]. IEEE Journal on Selected Areas in Communications, 2001, 21(8):1346–1357.
60. Ma D T, Wei J B, Zhuang Z W. Space laser communication and application [J]. Semiconductor Optoelectronic, 2003, 24(2): 139–144.
61. Ke X Z, Wu J L. Principle and application of wireless optical coherent communication [M]. Beijing: Science Press,2021.
62. Haas S M, Shapiro J H, Tarokh V. Space-Time Codes for Wireless Optical Channels [C]. IEEE International Symposium on Information theory, 2001: 244.
63. Zhu X M, Kahn J M, Wang J. Mitigation of Turbulence-Induced Scintillation Noise in Free-Space Optical Links Using Temporal-Domain Detection Techniques [J]. IEEE Photonics Technology Letters, 2003, 15(4):623–629.

64. Zhu X M, Kahn J M. Free-Space Optical Communication Through Atmospheric Turbulence Channels [J]. IEEE Transactions on Communications, 2002, 50(8): 1293–1300.
65. Simon M K, Vilnrotter V A. Alamouti-Type Space–Time Coding for Free-Space Optical Communication With Direct Detection [J]. IEEE Transactions on Wireless Communications, 2005, 4(1):35–39.
66. Alqudah Y A, Kavehrad M. Orthogonal Spatial Coding in Indoor Wireless Optical Link Reducing Power and Bandwidth Requirements [C]. Proceedings of SPIE - The International Society for Optical Engineering, 2003, 5285:237–245.
67. Pan F, Ma J, Tan L Y, et al. Scintillation characterization of multiple transmitters for ground-tosatellite laser communication [C]. Proceedings of SPIE - The International Society for Optical Engineering, 2005, 5640:448–454.
68. Wilson S G, Brandt-Pearce M, Cao Q, et al. Optical MIMO Transmission Using Q-ary PPM for Atmospheric Channels [C]. IEEE Conference on Signals, Systems & Computers, 2003: 1090–1094.
69. Brandt-Pearce M, Wilson S, Cao Q L, et al. Code Design for Optical MIMO Systems Over Fading Channels [C]. Proceedings of the 38th Asilomar conference on signals, systems & computers 2004: 871–875.
70. Simon M K, Vilnrotter V A. Alamouti-type Space-time Coding for Free-space Optical Communication with Direct Detection [J]. IEEE Transactions on Wireless Communications, 2005, 3(1):35–39.
71. Anguita J A, Neifeld M A, Vasic B V. Multi-Beam Space-Time Coded Systems for Optical Atmospheric Channels [C]. The International Society for Optical Engineering, SPIE, 2006, 6304: 332–340.
72. Anguita J A, Neifeld M A, Vasic B V. Spatial Correlation and Irradiance Statistics in A Multiple-beam Terrestrial Free-space Optical Communication Link [J]. Applied Optics, 2007, 46(26): 6561–6571.
73. Navidpour S M, Uysal M, and Kavehrad M. BER Performance of Free-Space Optical Transmission with Spatial Diversity [J]. IEEE Transactions on Wireless Communications, 2007, 6(8): 2813–2821.
74. Cao Q, Brandt-Pearce M, Wilson S G. Free Space Optical MIMO System Using PPM Modulation and a Single Optical Amplifier [C]. International Conference on Communications & Networking in China, IEEE, 2007:508–512.
75. Garcia-Zambrana A. Error Rate Performance for STBC in Free-space Optical Communications through Strong Atmospheric Turbulence [J]. IEEE Communications Letters, 2007, 11(5): 390–392.
76. Cvijetic N, Wilson S G, Brandt-Pearce M. Performance Bounds for Free-Space Optical MIMO Systems with APD Receivers in Atmospheric Turbulence [J]. IEEE Journal on Selected Areas in Communications, 2008, 26(3):3–12.
77. Chakraborty K, Dey S, Franceschetti M. Outage Capacity of MIMO Poisson Fading Channels [J]. IEEE transactions on information theory, 2008, 54(11):4887–4907.
78. Letzepis N, Holland I, Cowley W. The Gaussian Free Space Optical MIMO Channel with Q-ary Pluse Position Modulation [J]. IEEE Transactions on Wireless Communications, 2008, 7(5): 1744–1753.
79. Safari M, Uysal M. Do We Really Need OSTBCs for Free-space Optical Communication with Direct Detection [J]. IEEE Transactions on Wireless Communications, 2008, 7(11): 4445–4448.
80. Zaidi S R, Hafeez M. Cross Layer Design for Orthogonal Space Time block Coded Optical MIMO Systems [C]. Ifip International Conference on Wireless & Optical Communications Networks, IEEE, 2008:1–5.
81. Ntogari G, Kamalakis T, Sphicopoulos T. Performance Analysis of Space Time Block Coding Techniques for Indoor Optical Wireless Systems [J]. IEEE Journal on Selected Areas in Communications, 2009, 27(9):1545–1552.

82. Bayaki E, Schober R. On Space-time Coding for Free-space Optical Systems [J]. IEEE Transactions on Communications, 2010, 58(1):58–62.
83. García-Zambrana A, Castillo-Vázquez C, Castillo-Vázquez B. Space-time Trellis Coding with Transmit Laser Selection for FSO Links over Strong Atmospheric Turbulence Channels [J]. Optics Express, 2010, 18(6): 5356–5366.
84. Qiang X W, Wang T L, Wu N Q. Multi-beam laser atmospheric transmission [J]. Optoelectronics Technology, 1999, 19(3):167–172.
85. Liu W H, Wu J. The influence of atmospheric channel spacing and beam number on intensity fluctuation [J]. Optics and Optoelectronic Technology, 2003, 1(4): 15–19.
86. Liu W H, Wu J. The influence of multi-beam partially coherent light on intensity scintillation through strong turbulence [J]. Optical Engineering, 2004, 31(4):5–8.
87. Ma D T. Research on multi-beam emission and reception technology in atmospheric laser communication [D]. Changsha: National University of Defense Technology, 2004.
88. Ma D T, Hua W L, Wei J B. Temperature field analysis of laser source in multi-beam atmospheric laser communication [J]. Optoelectronic Laser, 2004, 15(12):1449–1451.
89. Ma D T, Wei J B, Zhuang Z W. Multi-beam transmission performance analysis and channel modeling in atmospheric laser communication [J]. Optics Journal, 2004, 24(8):1020–1028.
90. Chen G, Dong Z R, Gen J X, et al. 155 / 622 Mb/s multi-emitter laser communication system [J]. China Laser, 2004, 31(5):583–587.
91. Deng D Z, Guo H F. Multi-transmitter applications in atmospheric laser communications [J]. Optics and Optoelectronics Technology, 2005, 3(4):13–15.
92. Wu C J, Pan J M, Yu R J. Application of optimal selection diversity in atmospheric optical communication [J]. Infrared and Laser Engineering, 2005, 34(1):114–130.
93. He J, Hu Y J. Performance analysis of wireless optical MIMO system based on SVD algorithm [J]. Journal of Anhui University (Natural Science Edition), 2007, 31(3):34–37.
94. Deng Z, Li N. Diversity detection technology in laser communication [J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(3):280–284.
95. Tian Z R, Zhou H B. Structure and Performance of Space-Time Block Code [C]. International Conference on Wireless Communications, Networking and Mobile Computing, 2006.
96. Chen J W. Research on channel and performance of laser atmospheric transmission [D]. Wuhan: Huazhong University of Science and Technology, 2007.
97. Deng T P. Research on key technologies of free space optical communication system [D]. Wuhan: Huazhong University of Science and Technology, 2007.
98. Zhang H L. Research on Space-Time Coding Based on Transmit Diversity in Atmospheric Laser Communication [D]. Changchun: Changchun University of Science and Technology, 2008.
99. Chen C Y. Research on atmospheric influence mechanism and suppression technology in Optical Wireless Communication [D]. Changchun: Changchun University of Science and Technology, 2009.
100. Wang H Q, Ke X Z. MIMO atmospheric laser communication and its key technologies [J]. Laser Magazine, 2008, 29(3):52–54.
101. Simon M K, Alouini M S. Digital Communication over Fading Channels: A Unified Approach to Performance Analysis [M]. John Wiley &Sons, 2000:31–97.
102. Naguib A F, Calderbank R. Space-Time Coding and Signal Processing for High Data Rate Wireless Communications [M]. Springer US, 2002.
103. Zhao Y N, Zhang L L, Wu W L. The development and application of MIMO technology [J]. Telecommunication Technology. 2005,48(1): 7–11.
104. Ren L G, Song M, Chi S N, et al. MIMO technology in mobile communication [J]. Technology of Telecommunications, 2004(1):42–45.
105. Zhang X D, Bao Z. Communication signal processing [M]. Beijing: National Defense Industry Press, 2000.
106. Gesbert D, Akhtar J. Breaking the Barriers of Shannon's Capacity: An Overview of MIMO Wireless System[J]. Telenor's Journal: Telektronikk. 2002:1–9.

107. Telatar E. Capacity of Multi-antenna Gaussian Channels [J]. European Transactions on Telecommunications, 1999, 10(6):585-595.
108. Vucetic B, Yuan J. Space-time coding [M]. John Wiley & Sons, 2003.
109. Alamouti S M. A Simple Transmit Diversity Technique for Wireless Communications [J]. IEEE Journal on Selected Areas in Communications, 1998, 16(8):1451–1458.
110. Koike-Akino T, Popovski P, Tarokh V. Optimized Constellations for Two-way Wireless Relaying With Physical Network Coding [J]. IEEE Journal on Selected Areas in Communications, 2009, 27(5):773–787.
111. Tarokh V, Jafarkhani H, Calderbank A R. Space-time Block Codes from Orthogonal Designs [J]. IEEE Transactions on Information Theory, 1999, 45(5):1456–1467.
112. Larsson E G, Stoica P. Space-time Block Coding for Wireless Communications [M]. Cambridge: Cambridge University Press, 2003.
113. Foschini G J. Layered Space-time Architecture for Wireless Communication in A Fading Environment When Using Multi-element Antennas [J]. Bell Labs Technical Journal, 1996, 1(2):41–59.
114. Foschini G J, Chizhik D, Gans M J, et al. Analysis and Performance of Some Basic Space-Time Architectures [J]. IEEE Journal on Selected Areas in Communications. 2003, 21(3): 303–323.
115. Yang X. Comparative study of MIMO space-time coding techniques [J]. Information Technology, 2008, 32(6):103–105.
116. Tarokh V, Seshadri N, Calderbank A R. Space-time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction [J]. IEEE Transactions Information Theory, 1998, 44(2): 744–765.
117. Chen G, Cai Y M, Chen G T, et al. Discussion on some problems of space laser communication technology [J]. Infrared and Laser Engineering, 2000, 29(3): 44–48.
118. Mecherle G S, Henderson R J. Homodyne PSK Receivers with Laser Diode Sources [C]. Proceedings of SPIE-The International Society for Optical Engineering, 1991, 1417: 99-107.
119. Huang H X, Wu S L. DOA and Doppler frequency estimation algorithm for signal sources based on uniform circular array [J]. Electronic Journal, 2001, 29(5):619–624.
120. Zhao L. Research on DPPM modulation and demodulation of atmospheric laser communication [D]. Xi'an: Xi'an University of Technology, 2006.
121. An Y Q, Li W T. Optical Wireless Communication technology and application analysis [J]. Power System Communication, 2005, 26(150),4-6
122. Chen S Y, Han Y N. Overview of Optical Wireless Communication technology [J]. Guangxi Communication Technology, 2001, (1),39–41
123. Xu L L. Research on MIMO channel modeling [D]. Nanjing: Nanjing University of Technology, 2009.
124. Liu X. Space-time coding technology in MIMO [D]. Shanghai: Shanghai Jiao Tong University, 2006.
125. Ke X Z, Chen J, Deng L J. Research progress of space-time coding in Optical Wireless Communication (I) [J]. Infrared and Laser Engineering, 2013, 52(07):1882–1889
126. Ke X Z, Yuan L, Li F. Research progress of space-time coding in Optical Wireless Communication (II) [J]. Infrared and Laser Engineering, 2013, 42(08):2137–2145.
127. Ke X Z, Chen J, Li Z. Research progress of space-time coding in Optical Wireless Communication (III) [J]. Infrared and Laser Engineering, 2013, 42(09):2496–2504.
128. Ke X Z, Chen J, Chen D. Research progress of space-time coding in Optical Wireless Communication (IV) [J]. Infrared and Laser Engineering, 2013, 42(10):2765–2771.

Space–Time Coding

20

Contents

Introduction 632
Overview of Space–Time Coding Research 632
Rise of Optical Wireless MIMO 632
Domestic Research on Optical Wireless MIMO 635
Multi-Aperture Transmitting/Receiving Experiment 635
Multibeam-Emission System Composition and Experiment 635
Multi-aperture Receiving-System Composition and Related Experiments 637
Layered Space–Time Codes 640
Layered Space–Time Coding Models 640
Layered Space–Time Codes 641
Horizontal Layered Space–Time Coding 642
Diagonal Layered Space–Time Codes 645
Helical Layered Space–Time Coding 649
Vertical Layered Space–Time Coding 652
Space–Time Codes 658
Orthogonal Space–Time Coding 658
Space–Time Trellis Coding 661
Differential Space–Time Coding 664
Unitary Space–Time Coding 669
Hybrid Space–Time Coding 672
Summary 674
Concatenated Code 674
Cascaded Orthogonal Space–Time Coding 675
Cascaded Layered Space–Time Codes 679
References 686

Abstract

Multiple-input multiple-output (MIMO) technology uses multiple antennas at the sending and receiving ends, which can suppress channel fading, reduce the bit-error rate, and improve spectrum utilization. It makes space a resource that can be used to improve communication performance. In wireless laser communication, space–time coding combines the advantages of space and time diversity

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_20

and simultaneously provides diversity and coding gains. Space–time coding can not only improve the channel capacity of communication systems but also suppress channel fading.

Keywords

Optical wireless communication · Hierarchical space–time code · Space–time code · Cascade code

Introduction

Space–time coding is a signal coding technique that can obtain high data transmission rate, which is a two-dimensional processing method that combines spatial and temporal transmission signals. With the continuous development of wireless communication technology, space–time coding has been studied by more and more scholars due to its ability to use multiple antennas at both the transmitter and receiver ends to effectively counteract channel fading, reduce the bit error rate, improve spectrum utilization [1–4], and improve system performance.

Overview of Space–Time Coding Research

Rise of Optical Wireless MIMO

Since 1960, extensive research has been conducted on wireless laser communication [5–9]. International organizations engaged in optical wireless communication research mainly include the European Space Agency (ESA), Communications Research Office (CTL) of the Ministry of Posts of Japan, Japan's National Space Development Agency (NASDA), United States National Aeronautics and Space Administration (NASA), and United States Air Force.

In 1980, the Massachusetts Institute of Technology (MIT) Lincoln Laboratory conducted an end-to-end satellite-communication demonstration experiment with a high bit rate using the heterodyne method [9]. In 1985, the US Air Force developed a half-duplex optical-link connection experiment between a ground station and the laser intersatellite transmission experiment (LITE) system [7]. In the late 1980s, the European Space Agency established a laser communication link between two satellites and verified the technology of an intersatellite laser communication unit [5, 8].

In 1982, Nakai studied coding and modulation technology in an optical channel, laying the foundation for the development of optical communication [10]. Later, Reed–Solomon (RS) codes [11], trellis-coded modulation (TCM) codes [12], Bose–Chaudhuri–Hocquenghem (BCH) codes [13], turbo codes [14, 15], low-density parity-check (LDPC) codes [16], interleaving technology [17], and concatenated codes [18, 19] were applied to wireless laser communication.

However, it is still difficult to meet the requirements of large-capacity, high-speed, and long-distance communication.

In 1995, Japan successfully conducted a satellite–ground optical-link communication experiment between a ground station and Engineering Test Satellite VI (ETS-VI). In the same year, the US Strategic Missile Defense Organization developed an all-weather tracking and scanning test system. In December 1996, Thermo Trex conducted an optical communication experiment between a ground station and an aircraft. Foreign manufacturers engaged in atmospheric laser communication include Air Fiber, Light Pointe, Terabeam, and Canon.

In 1997, Lesh and Wilson et al. of Caltech first proposed the concept of multibeam emission [20] to suppress turbulence effects in atmospheric laser transmissions and effectively compensate atmospheric channels.

Since the concept of multibeam emission was proposed, the MIT Lincoln Laboratory and the Jet Propulsion Laboratory (JPL) of the United States have experimented with multibeam atmospheric transmissions and obtained numerous experimental results [21–24]. From October 1995 to May 1996, the United States and Japan successfully verified the effectiveness of multibeam emission technology in suppressing atmospheric-turbulence effects [25]. Kim et al. conducted an atmospheric-scintillation measurement experiment [21] and found that the fluctuation in the received light intensity caused by atmospheric turbulence decreased with an increase in the number of transmitted beams. A theoretical model of multibeam propagation in a turbulent atmosphere was presented in [26, 27].

In 1996, assuming that the input-signal fading of each small-aperture receiver was independent, Ibrahim [28] studied the performance of wireless laser communication using reception-diversity technology. In 2000, Bell Labs successfully developed a prototype system for atmospheric laser communication using multibeam transmission technology [29]. In 2002, the JPL conducted an optical-transmission experiment at a distance of 46.8 km using multiple beams [30].

The JPL has studied a combination of output signals and conducted field experiments [31–33]. In 2001, the JPL's Communication Engineering Research Office and Stanford University's Laser-Communication Research Group conducted in-depth research on diversity reception and optical-signal detection [34].

In the same year, Haas et al. [35] of the Massachusetts Institute of Technology proposed space–time coding criteria for atmospheric laser communication and provided their error-code performance limits. However, this theory is only suitable for heterodyne detection. Heterodyne detection requires that the frequencies of the receiver's local oscillation laser and the emission laser be very stable, and that the frequency stability must be at least 10–11 orders of magnitude, which affects the practicality of the process [36]. Reference [37] also provides space–time coding criteria and error-code performance limits; however, it remains difficult to apply this method to direct-detection atmospheric laser communication systems.

In 2002, Zhu et al. [38–41] applied the time-domain detection technology, likelihood-ratio detection theory, and diversity-reception technology in radio-frequency (RF) communication to atmospheric laser communication, and studied their effectiveness in suppressing turbulence effects. In 2003, Simon et al. improved

the traditional Alamouti code [40] and proposed a method to realize space–time coding in atmospheric laser communication using intensity modulation/direct detection (IM/DD).

In the same year, Alqudah et al. proposed orthogonal space–time coding for indoor optical wireless channels [41]. In 2004, Feng et al. [42] analyzed the light-intensity fluctuation characteristics of a satellite–ground multibeam transmission link, and Wilson et al. analyzed the bit-error performance of MIMO optical communication using quantized pulse-position modulation (QPPM) technology [43]. Under the same conditions, Brandt-Pearce et al. [44] designed a coding scheme with equal Hamming weights using PPM modulation technology and analyzed the error-code performance limit.

In 2005, Simon et al. proposed a space–time coding scheme suitable for IM/DD [45]; however, this scheme was only applicable to on–off keying (OOK) modulation and 2-PPM modulation. In 2006, Angulta et al. [46] analyzed the space–time coding mode for communication systems using multibeam transmissions.

In 2007, Angulta et al. [47] found that the ability to reduce fluctuations in light intensity was related to the spatial correlation of the beams, transmission distance, turbulence conditions, transmitter spacing, and receiver aperture. Navidpour et al. [48] of Pennsylvania State University found that the spatial-diversity performance was seriously affected by the spatial correlation of each subchannel. In the same year, Cao et al. [49] proposed a MIMO system for atmospheric laser communication suitable for PPM modulation. In the same year, García-Zambrana analyzed the bit-error performance of the space–time block codes proposed by Simon in strong turbulence channels [50].

In 2008, Neda et al. [51] analyzed a ground optical communication link during multibeam transmissions and multiple avalanche-photodiode (APD) detection and reception. For IM/DD optical communication, reference [52] analyzed the MIMO interrupt capacity in a Poisson fading channel with thermal noise and limited transmission power. Letzepis et al. used PPM modulation to study the performance of MIMO systems in free-space optical (FSO) communication under a Gaussian mechanism [53].

Safari et al. [54] studied the performance of repetitive codes and space–time block codes using OOK modulation, while Raza et al. [55] proposed a cross-coding scheme by combining the media access control (MAC)-layer and physical-layer coding (orthogonal space–time coding), and used training sequences to estimate the channel parameters.

In 2009, Ntogari et al. [56] studied the performance of space–time coding in a scattering-rich indoor environment, in combination with IM/DD technology. In 2010, Bayaki et al. [57] introduced a closed type with the probability of pairing errors in the space–time coding of free-space optical communication, when receiving two lasers and any detector for IM/DD. In the same year, García-Zambrana et al. [58] proposed a space–time trellis code for a strong turbulence channel and analyzed its bit-error performance.

It is necessary to further study the structure, design, performance, and scope of space–time coding applications in IM/DD atmospheric laser communication.

Domestic Research on Optical Wireless MIMO

Research on multibeam transmitting and receiving technologies has been gradually conducted in China. In 1999, the Northwest Nuclear Technology Research Institute studied the transmission of a multibeam laser in the atmosphere. The results showed that the received light intensity of a multibeam transmission was twice as high as that of a single-beam transmission [59]. In 2003–2004, the University of Electronic Science and Technology of China studied the theory of multibeam transmissions and showed that multiple beams could effectively suppress the atmospheric-turbulence effect [60, 61].

In 2004, the National University of Defense Technology [62–64] studied a multibeam-transmission channel model, space–time coding, and the temperature field of multibeam light sources. The Shanghai Institute of Optics and Precision Machinery, Chinese Academy of Sciences, implemented a 155/622 Mbit/s multi-transmitter laser communication system [65].

In 2005, the Huazhong Optoelectronic Technology Research Institute [66], Yanshan University [67], and others studied diversity technology in atmospheric laser communication and analyzed the system performance. Anhui University [68], Beijing University of Aeronautics and Astronautics [69], and other organizations have published relevant reports.

In 2006, Wuhan University used quadrature amplitude modulation (QAM) to apply the space–time coding in RF communication to optical wireless communication, and verified the performance of space–time block coding through simulation experiments [70]. Since 2007, Xi'an University of Technology has investigated space–time coding technology in atmospheric laser communication and conducted optical MIMO communication experiments. Huazhong University of Science and Technology [71, 72] and Changchun University of Technology [73, 74] have also studied space diversity, time diversity, optical MIMO technology, and space–time coding in optical wireless communication.

Multi-Aperture Transmitting/Receiving Experiment

Multibeam-Emission System Composition and Experiment

Figure 20.1 shows the composition of an experimental multibeam emission system. The system consists of two parts: transmitting and receiving. In this experiment, the transmitting end uses two lasers with the same wavelength. The laser uses 532-nm green light. The receiving end uses a single aperture. The receiving lens has a 0.1-m diameter. The laser is transmitted through a Cassegrain telescope into the atmosphere.

At the receiving end, the signal is gathered into a beam of light through the receiving lens, and a detector is used for detection. The detector is connected to a data-acquisition card on a computer. The measured data were used to calculate the light-intensity variance. The experiment was conducted at 9:00 p.m. on November

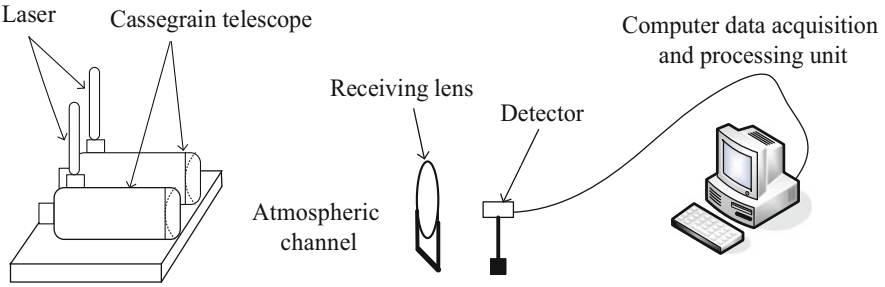


Fig. 20.1 Composition of the multibeam-emission experimental system

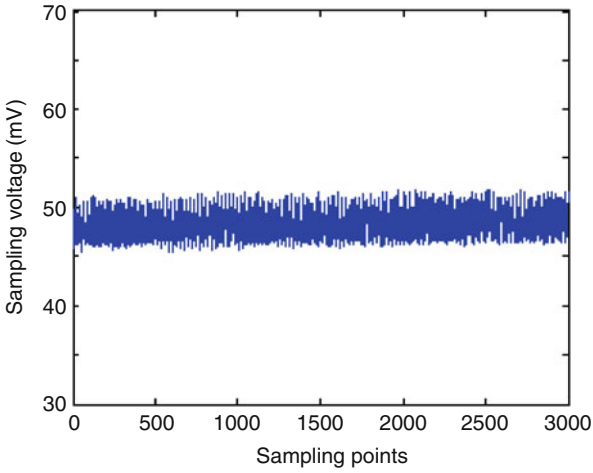


Fig. 20.2 Measured voltage values of a single transmitter and single receiver

24, 2011, between the sixth and fifth floors of Xi'an University of Technology. It was sunny and breezy during the experiment.

The light-intensity data of single-beam and dual-beam emissions were collected in the experiment to compare the performance of single-emission and multiple-emission systems. Figure 20.2 shows the sampling voltages measured when transmitting a single beam and receiving with a single aperture. The sampling frequency was set at 2 MHz. Three-thousand data points were used in the experiment. The figure shows the discrete points. It can be seen from the figure that most of the data were concentrated near 47, and a small number of data points were slightly off-center, which was caused by light-intensity flickering.

Figure 20.3 shows the sampling voltage when transmitting with two light beams and receiving with a single lens. It can be seen from the figure that, compared with a single-beam transmission, the voltage fluctuation when transmitting with two light beams is reduced and relatively stable.

According to the theoretical calculations, the variance values of the light intensity of a single transmitter with a single receiver and a double transmitter with a single

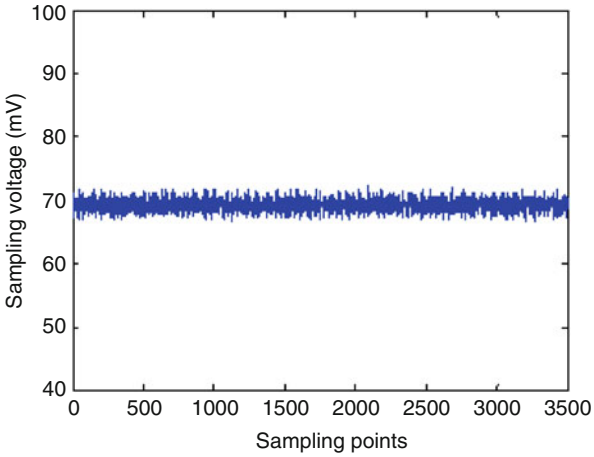


Fig. 20.3 Measured voltage values of double-beam transmitting and single-aperture receiving

receiver were 3.4001 and 2.3800, respectively, and the variance values of the data measured in the experiment were 1.8815 and 0.8852, respectively. Theoretical and experimental measurement data show that the variance of the light intensity of the double-beam emission is smaller than that of the single-beam emission. Thus, a multibeam emission system can reduce the light-intensity fluctuation and improve the communication-system performance.

Multi-aperture Receiving-System Composition and Related Experiments

Figure 20.4 shows the structural diagram of the multi-aperture receiving system. The transmitting end is mainly composed of a Cassegrain telescope and a laser, and the receiving end is composed of a receiving lens with three apertures, beam-convergence lens array, detector, and computer data-acquisition and processing unit. The spacing between the three lenses is 0.1 m, and the diameter of the lens is 0.1 m.

When the laser beam at the transmitting end reaches the surface of the receiving lens and passes through it, three beams of light appear. It is necessary to use a lens and a reflector array that converge the light beam to combine the three beams into one beam. However, the optical path is relatively complex. The key point is to maintain the stability of the transmitting beam; otherwise, the beams are difficult to combine.

The experiment was conducted at 9:00 p.m. on October 25, 2011, when the weather was cloudy. The laser used in the experiment was a green light with a 532-nm wavelength. The data-acquisition card at the receiver had a 40-MHz sampling frequency. When the beam passed through the three-aperture receiving lens, it was received by the detector and collected using a computer. During the experiment,

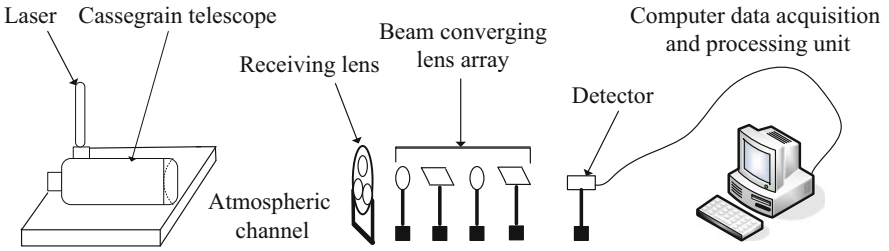
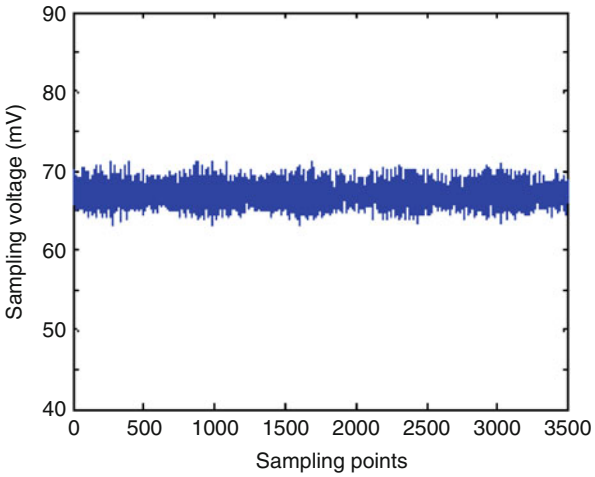


Fig. 20.4 Structural diagram of a multi-aperture receiving system

Fig. 20.5 Sampling voltage for single-aperture reception



two groups of data received by one, two, and three apertures were measured. Figures 20.5, 20.6, and 20.7 show sample voltage values when single-aperture, dual-aperture, and triple-aperture reception were adopted, respectively.

Comparing the single-, dual-, and triple-aperture reception, it can be seen that with an increase in the number of apertures, the measured data jitter decreases, and the overall signal is relatively stable. The variance was calculated based on the data measured in this experiment.

Two groups of data were calculated in each case. The measured variances of the single-aperture reception were 5.9763 and 5.717, respectively, and the theoretically calculated variance was 4.8638. The variances of the experimental data for dual-aperture reception were 2.8183 and 2.8181, respectively, and the theoretical data was 2.0372. When the triple-aperture receiver was used, the two groups of variances calculated from the experimental data were 0.1042 and 0.1050, respectively, while the variance calculated from the theoretical data was 0.5998.

In general, it can be concluded that the light-intensity variance decreases with an increase in the number of apertures, indicating that multi-aperture reception can more

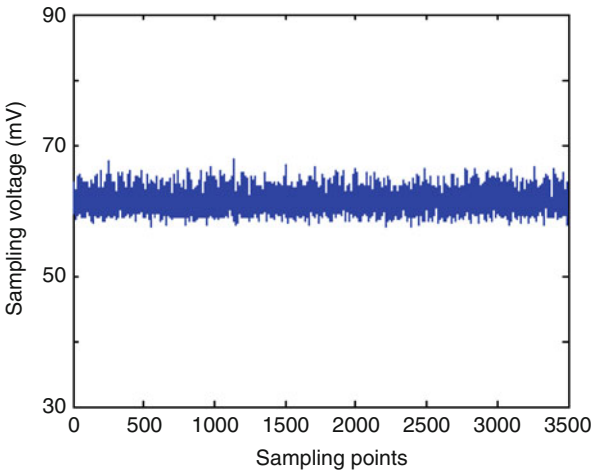


Fig. 20.6 Sampling voltage for dual-aperture reception

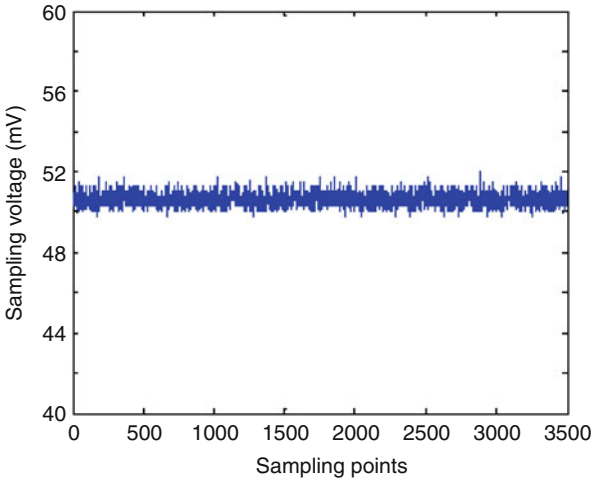


Fig. 20.7 Sampling voltage for triple-aperture reception

effectively suppress the impact of turbulence and improve the laser-transmission performance.

We conducted experiments on multibeam transmission and multi-aperture reception, and calculated the light-intensity variance of the data measured in the experiment. By comparing the light-intensity fluctuation variance of different numbers of beams when transmitting and the light-intensity variance of different numbers of receiving apertures when receiving, we can conclude that the multibeam-transmission system and multi-aperture reception system could suppress the impact of turbulence on laser transmissions more effectively than single-beam transmission and single-aperture reception systems.

Layered Space–Time Codes

Layered space–time codes (LSTCs) [1, 75] are based on the multiple-input multiple-output (MIMO) transmission mode proposed by Bell Laboratories. It is mainly based on space-division multiplexing to improve the spectral efficiency of the system. The Bell Laboratories layered space–time (BLAST) code [72] can be coded in both spatial and temporal dimensions. It decomposes the input information bitstream into several bitstreams, and independently encodes, modulates, and maps them to multiple transmission antennas.

At the receiving end, special processing technology is used to separate the signals that arrive at the receiving antenna and send them to the corresponding decoder [77]. This technology combines a one-dimensional coding method with spatial dimension transmission. Thus, the computational complexity of the decoding end only increases linearly with an increase in the number of antennas, unlike the exponential growth of the two-dimensional space–time code.

Layered space–time coding requires that the number of receiving antennas should not be less than the number of transmitting antennas, which means that multiple antennas should be installed at the receiving end. However, it is relatively difficult for RF communication systems to install multiple antennas on mobile terminals, which will undoubtedly limit the application of this technology in RF communication. Multiple receiving antennas are no longer difficult in optical wireless communication.

Layered Space–Time Coding Models

Layered space–time codes include horizontal (H-BLAST), diagonal (D-BLAST), vertical (V-BLAST), and spiral layered space–time codes (S-BLAST).

We combined the OOK and PPM intensity-modulation methods to construct a layered space–time coding scheme [78] in a optical wireless channel. The transmission model is illustrated in Fig. 20.8. The basic idea is to divide high-speed data services into several low-speed data services, which are sent to M identical channel encoders for coding. The channel coding is either block or convolutional. After channel coding, it is layered space–time coded and then transmitted by M antennas after modulation.

The receiving model is illustrated in Fig. 20.9. At the receiving end, N antennas are used for diversity reception. The layered feedback interference is eliminated by the linear decision-feedback equalizer and sent to the channel decoder after layered space–time decoding. The channel must be estimated because the channel-transmission characteristics are used in the decoding process.

The reception model of the layered space–time coding shows that the received signal at each receiving aperture is the superposition of signals from different transmitting apertures. In a layered space–time structure, multiple transmit antennas are used to achieve high data rates, and multiple receive antennas are used to cancel the interference of multiple spatial signals.

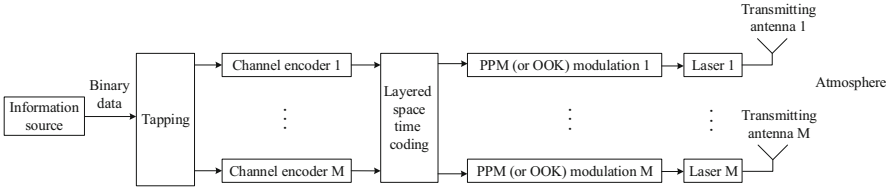


Fig. 20.8 Transmission model of layered space–time coding in atmospheric laser communication

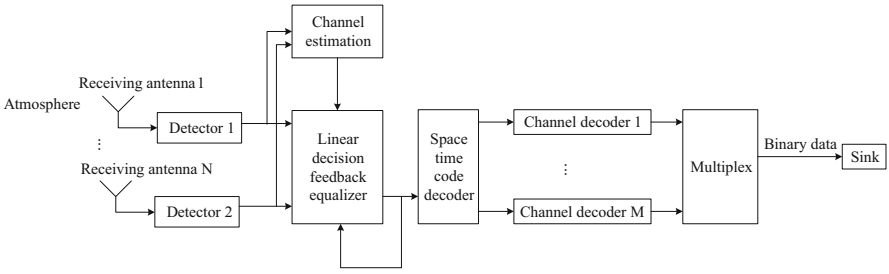


Fig. 20.9 Reception model of layered space–time coding in atmospheric laser communication

Layered Space–Time Codes

We used the horizontal-layered, spiral-layered, and diagonal-layered space–time codes as examples to analyze the performance of layered space–time codes. The simulation conditions are as follows:

1. Assuming that the total power E_s is unchanged, the value is 1 (normalized).
2. We assume that the channel characteristics and noise can be accurately estimated by the receiver.
3. The photoelectric-conversion efficiency is 0.6.
4. In weak turbulence, the atmospheric channel obeys a logarithmic normal distribution, and the scintillation index SI is 0.6. The atmospheric channel follows a gamma–gamma distribution in moderate–strong and strong turbulence; the scintillation indexes SI are 1 and 3, respectively.
5. The system adopts pulse-position modulation, taking 4-PPM modulation as an example.
6. 2×2, 2×4, 4×4, and 4×6 antennas are used as examples.

Zero-forcing (ZF) detection, minimum mean-squared error (MMSE) detection, maximum-likelihood (ML) detection, sequential interference-cancellation (SIC), and ordered sequential interference-cancellation (OSIC) algorithms were used to analyze the performance of the layered space–time codes.

Horizontal Layered Space–Time Coding

Figures 20.10 and 20.11 show the error-rate curves under the zero-forcing detection, minimum mean-squared error detection, and maximum-likelihood detection algorithms when SI = 0.6. The weak-turbulence atmospheric channel follows a lognormal distribution. The following results can be observed:

1. The detection algorithm with the best performance is the maximum-likelihood detection algorithm, followed by the minimum mean-squared error detection algorithm, and finally the zero-forcing detection algorithm. The simulation results were consistent with the theoretical analysis results.
2. The diversity gain is related to the number of detectors, as shown in Fig. 20.11.

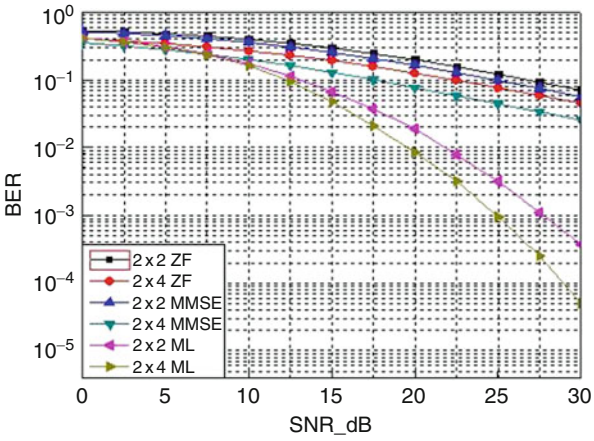


Fig. 20.10 Simulation diagram of 2 × 2 and 2 × 4 (SI = 0.6)

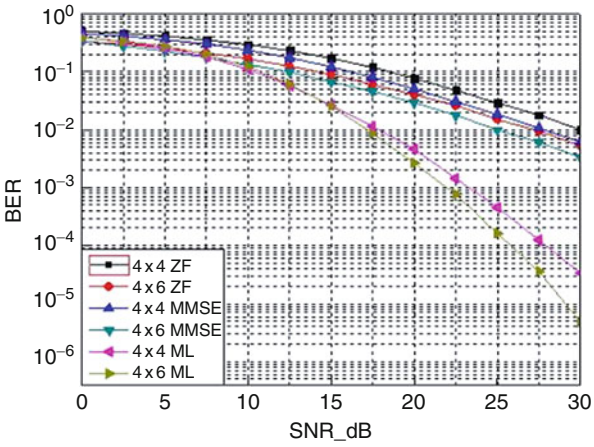


Fig. 20.11 Simulation diagram of 4 × 4 and 4 × 6 (SI = 0.6)

Figures 20.12 and 20.13 show the error-rate curves of the zero-forcing detection, minimum mean-squared error detection, and maximum-likelihood detection algorithms when SI = 1. The atmospheric channel follows a strong-turbulence gamma–gamma distribution. It can be seen that atmospheric turbulence affects the system performance to a certain extent; however, the bit-error rate (BER) characteristics, similar to those in weak turbulence, can still be obtained when the number of antennas is increased.

Figures 20.14 and 20.15 show the bit-error rate characteristics of the system under strong turbulence when SI = 3. It can be observed that the bit-error rate of the system is affected by strong turbulence in the case of two-transmitter–multiple receivers and four-transmitter–multiple receivers. However, the bit-error rate of the system with

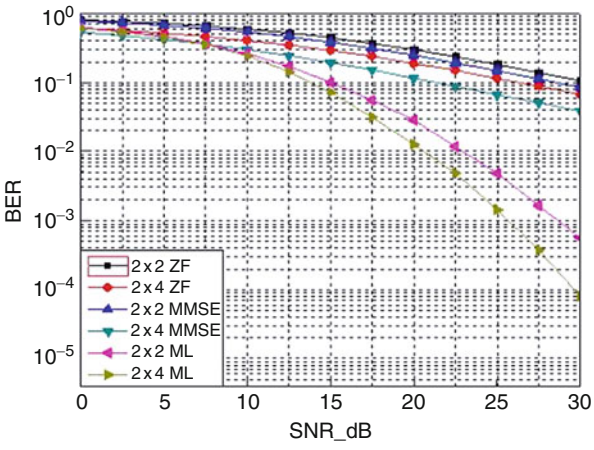


Fig. 20.12 Simulation diagram of 2 × 2 and 2 × 4 (SI = 1)

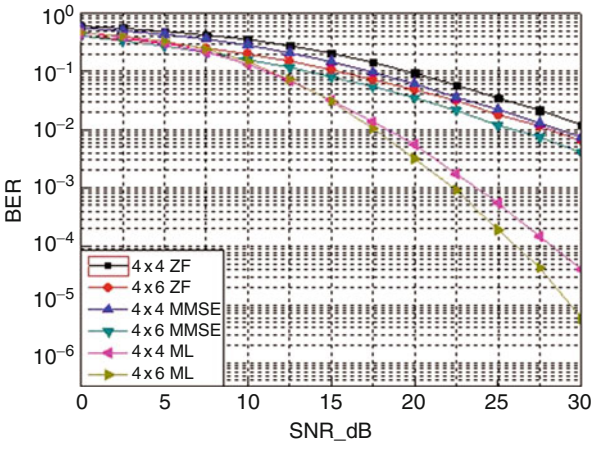


Fig. 20.13 Simulation diagram of 4 × 4 and 4 × 6 (SI = 1)

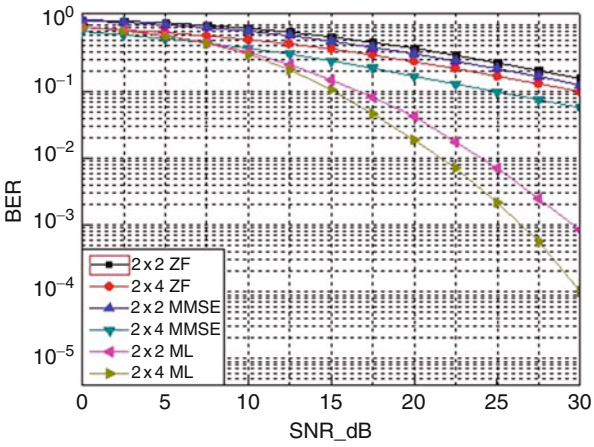


Fig. 20.14 Simulation diagram of 2×2 and 2×4 (SI = 3)

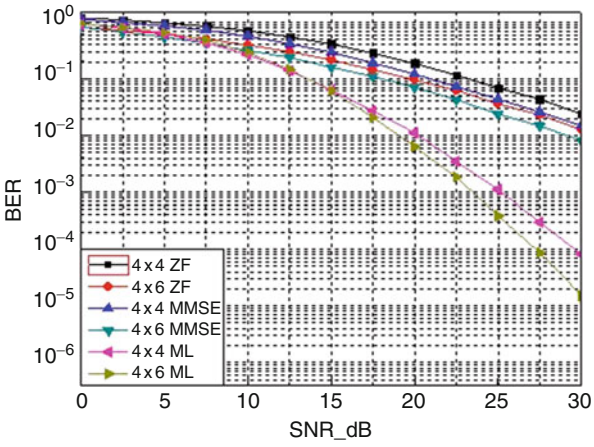


Fig. 20.15 Simulation diagram of 4×4 and 4×6 (SI = 3)

two transmitters and multiple receivers greatly increased, from nearly 10^{-6} to 10^{-4} at 30 dB, and the bit-error rate of the system with four transmitters and multiple receivers rose to 10^{-5}.

Horizontal layered space–time coding cannot effectively suppress the impact of atmospheric turbulence on the system, owing to its poor coding characteristics. Figure 20.16 compares the ZF, SIC-ZF, and OSIC-ZF detection algorithms in a horizontal layered space–time coding system. It can be seen that SIC-ZF and OSIC-ZF perform better than the ZF algorithm, which improves the performance of the horizontal layered space–time coding system to a certain extent.

Figure 20.17 compares the results of three detection algorithms, MMSE, SIC-MMSE, and OSIC-MMSE, in a horizontal layered space–time coding system.

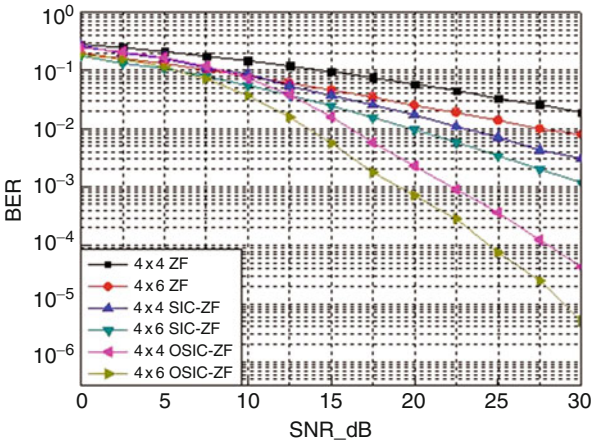


Fig. 20.16 Simulation of ZF, SIC-ZF, and OSIC-ZF algorithms

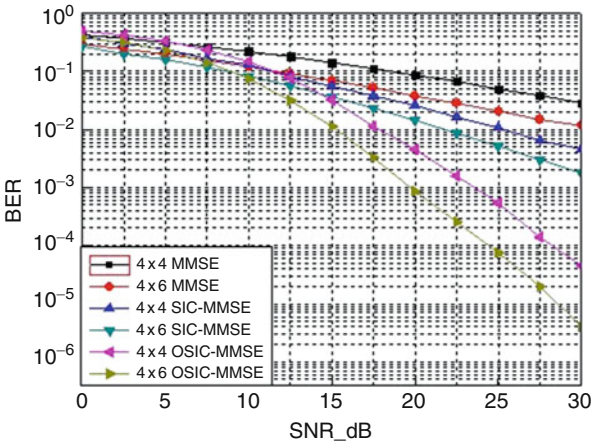


Fig. 20.17 Simulation of MMSE, SIC-MMSE, and OSIC-MMSE algorithms

It can be seen that the ordered successive interference-cancellation (OSIC) algorithm improves the performance of the MMSE detection algorithm, and the improvement extent is greater than that of the serial detection algorithm.

Diagonal Layered Space–Time Codes

Considering different receiving and transmitting antennas, noise levels, and atmospheric scintillation-factor parameters, the impact of diagonal layered space–time coding on the system bit-error rate is analyzed.

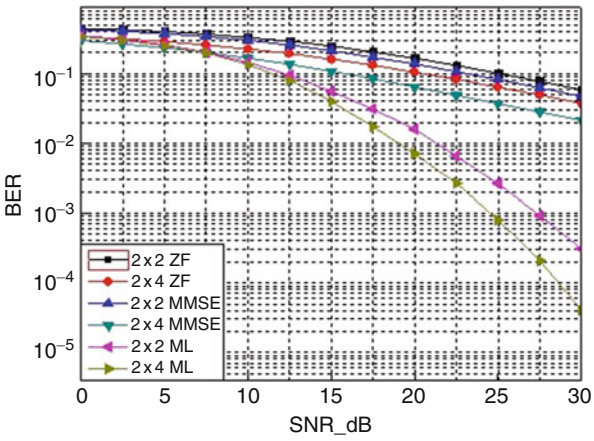


Fig. 20.18 Simulation diagram of 2 × 2 and 2 × 4 (SI = 0.6)

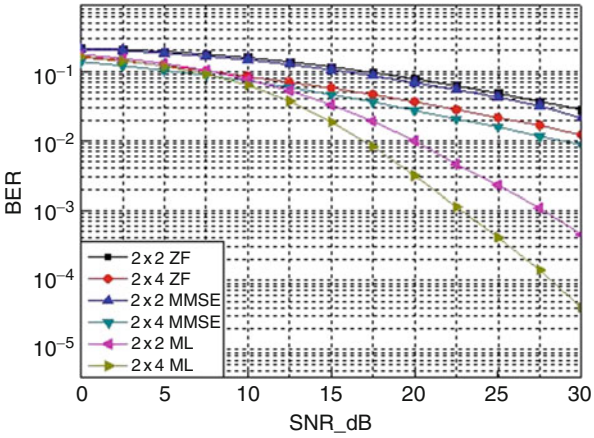


Fig. 20.19 Simulation diagram of 2 × 2 and 2 × 4 (SI = 1)

Comparing Figs. 20.18 and 20.19, when the scintillation factors SI are 0.6 and 1, respectively, in the case of two transmitters and multiple receivers, the bit-error rate of the system slightly increases in the case of medium turbulence with an increase in the scintillation factor. This indicates that the system cannot effectively suppress the impact of atmospheric turbulence in this case.

Comparing Figs. 20.20 and 20.21, in the case of four transmitters and multiple receivers, enhancing the turbulence intensity does not affect the bit-error rate of the system. Increasing the number of antennas under moderate and strong turbulence enhances the system's ability to suppress atmospheric turbulence. It can be seen that the D-BLAST system has a certain ability to suppress the scintillation effect caused by atmospheric turbulence.

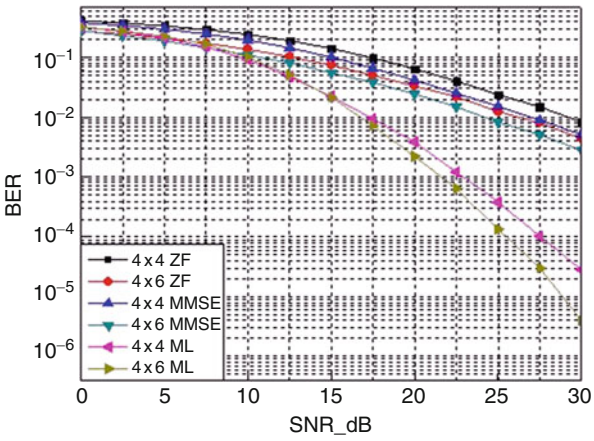


Fig. 20.20 Simulation diagram of 4 × 4 and 4 × 6 (SI = 0.6)

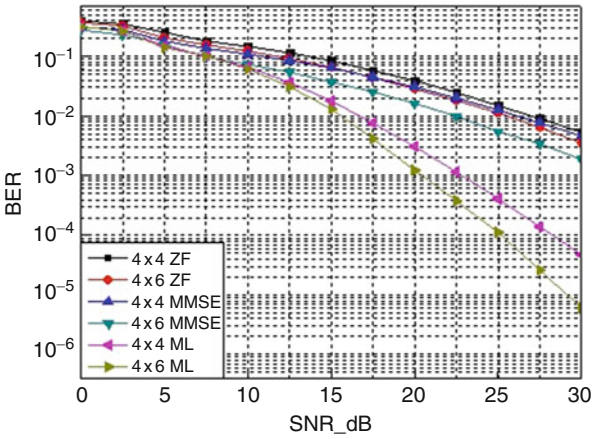


Fig. 20.21 Simulation diagram of 4 × 4 and 4 × 6 (SI = 1)

The performance of the D-BLAST coding system is also greatly affected by different detection algorithms. As shown in Fig. 20.21, when the scintillation index was SI = 1 and the signal-to-noise ratio (SNR) was approximately 30 dB, the numbers of antennas were 4 × 4 and 4 × 6. The ZF bit-error rates were 3×10^{-2} and 1×10^{-2}, respectively. The bit-error rates obtained using MMSE were 8×10^{-2} and 5×10^{-3}, respectively. The ML error rates were 1×10^{-4} and 4×10^{-6}. The performance of ML in the D-BLAST coding system was significantly better than that of MMSE and ZF.

Figures 20.22 and 20.23 show the BER and SNR curves of the system under strong turbulence and the scintillation index SI = 3. In the case of strong turbulence, the BER of the two transmitter–multiple receiver systems increased from nearly 10^{-6} to 10^{-5} at 30 dB. However, in the case of four transmitters and multiple

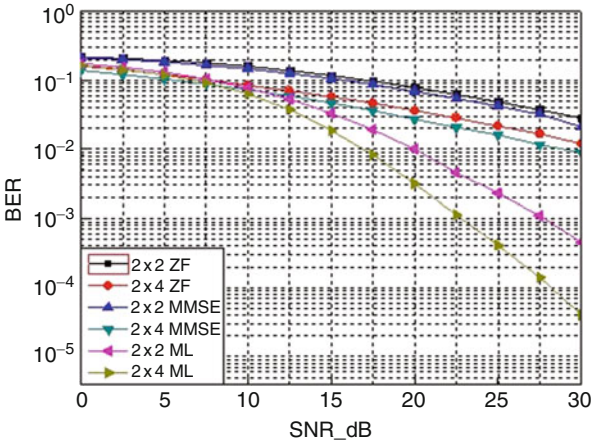


Fig. 20.22 Simulation diagram of 2 × 2 and 2 × 4 (SI = 3)

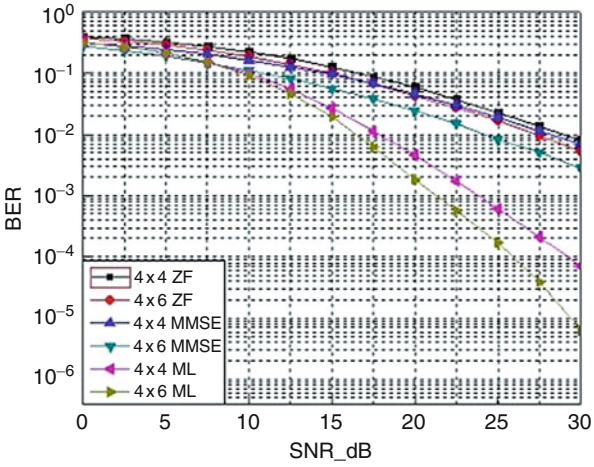


Fig. 20.23 Simulation diagram of 4 × 4 and 4 × 6 (SI = 3)

receivers, the bit-error rate of the system did not change much. Moreover, after increasing the number of antennas, the system restrained the impact of atmospheric turbulence to a certain extent.

As shown in Figs. 20.24 and 20.25, the performances of the SIC-ZF, OSIC-ZF, SIC-MMSE, and OSIC-MMSE detection algorithms were significantly better than those of ZF and MMSE, and the system's bit-error performance was significantly improved. The performance of OSIC was better than that of SIC alone. For example, when the BER is 10^{-2}, the SNR of the system using the OSIC-MMSE algorithm is about 15 dB higher than that of the system using the MMSE algorithm. The signal-to-noise ratio of the system using the SIC-MMSE algorithm improved by about 10 dB.

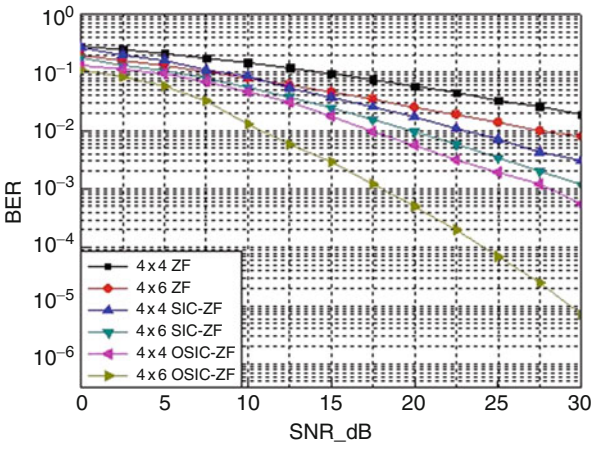


Fig. 20.24 Simulation of ZF, SIC-ZF, and OSIC-ZF

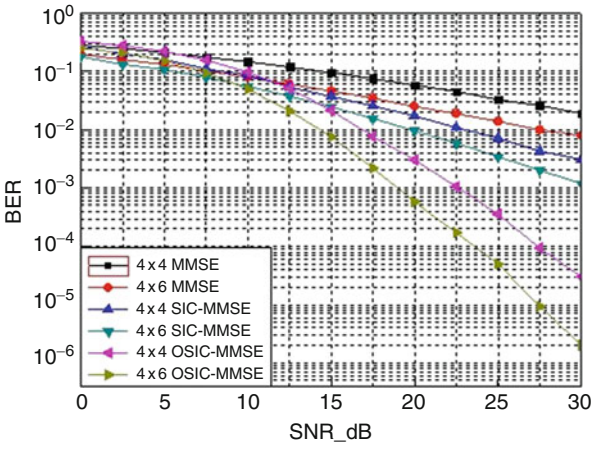


Fig. 20.25 Simulation of MMSE, SIC-MMSE, and OSIC-MMSE

Helical Layered Space–Time Coding

As shown in Fig. 20.26, when the SNR is 30 dB and the detection algorithm is ML, the bit-error rates for 2×2 and 2×4 are respectively 10^{-3} and 10^{-4}. When the signal-to-noise ratio is small, the performance gap between the ZF and MMSE systems is large, and the performance of the MMSE detection algorithm is better when the SNR is low.

As shown in Fig. 20.27, the maximum-likelihood detection algorithm was used in the T-BLAST coded system. When the SNR is 30 dB, the system's bit-error rate drops from 10^{-4} to near 10^{-6}. The performance of the maximum-likelihood detection algorithm was better than those of the MMSE and ZF detection algorithms.

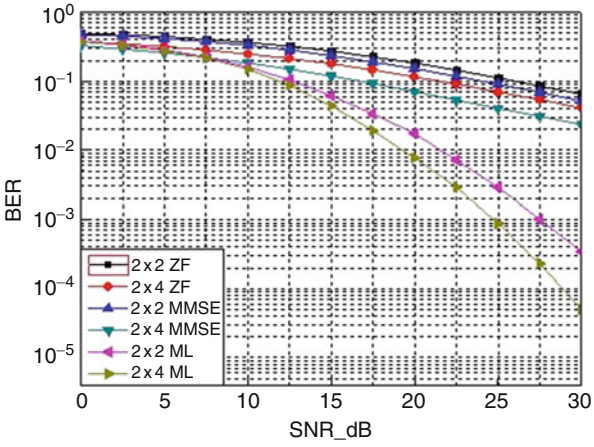


Fig. 20.26 Simulation diagram of 2×2 and 2×4 (SI = 0.6)

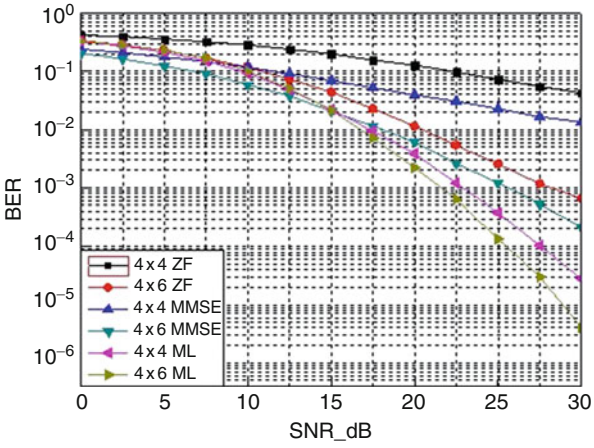


Fig. 20.27 Simulation diagram of 4×4 and 4×6 (SI = 0.6)

Comparing Figs. 20.27 and 20.28, the scintillation index SI is 0.6 and 1, respectively. To reduce the bit-error rate, the T-BLAST system must increase the number of transceiver antennas. The T-BLAST and D-BLAST FSO MIMO systems can well overcome the influence of medium and strong turbulence on the system (Fig. 20.29).

Figures 20.30 and 20.31 show the simulation curves of the bit-error rate and SNR of the system under strong turbulence. The scintillation index SI = 3 indicates that the system will be affected by strong turbulence in the atmosphere. The bit-error rate of the system with two transmitters and multiple receivers increases to a certain extent from nearly 10^{-6} to 10^{-4} at 30 dB. Increasing the number of transmitting and receiving antennas will reduce the influence of turbulence on the system performance to a certain extent.

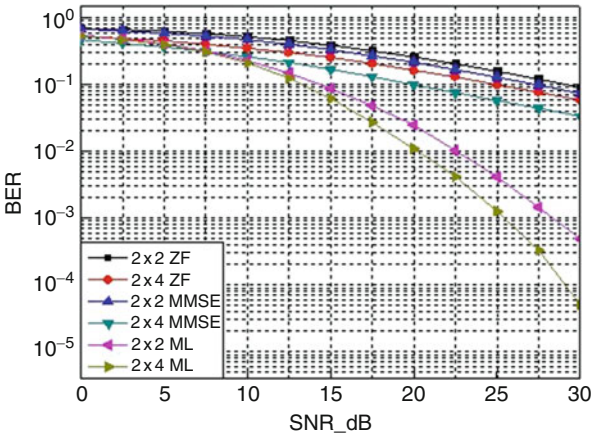


Fig. 20.28 Simulation diagram of 2×2 and 2×4 (SI = 1)

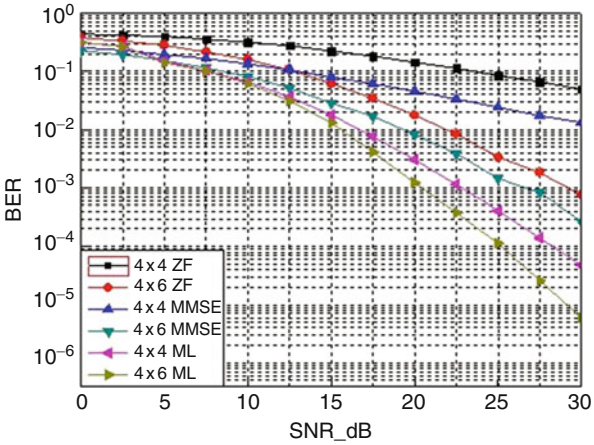


Fig. 20.29 Simulation diagram of 4×4 and 4×6 (SI = 1)

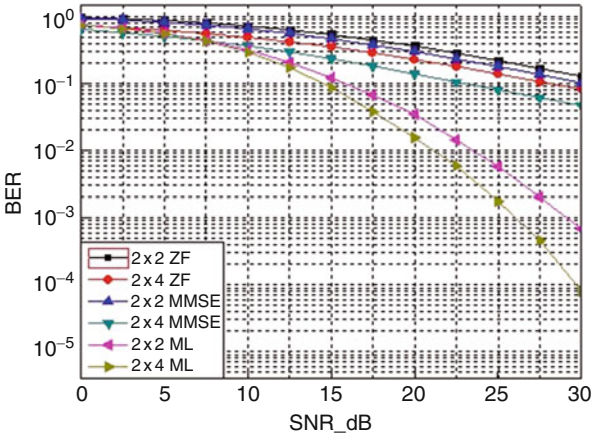


Fig. 20.30 Simulation diagram of 2×2 and 2×4 (SI = 3)

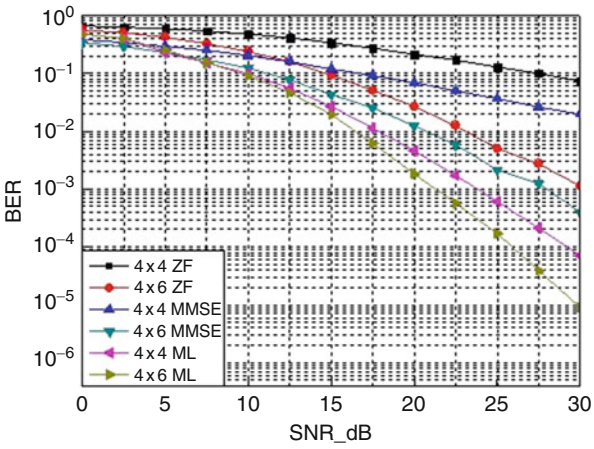


Fig. 20.31 Simulation diagram of 4×4 and 4×6 (SI = 3)

Figures 20.32 and 20.33 show the simulation results of the successive interference-cancellation (SIC) and ordered successive-interference cancellation (OSIC) algorithms using T-BLAST coding. It can be seen that the bit-error rate of the system using spiral layered space–time coding is also close to 10^{-6}. It also shows that the performance of the improved SIC-ZF, OSIC-ZF, SIC-MMSE, and OSIC-MMSE detection algorithms was significantly better than that of the ZF and MMSE detection algorithms, and their bit-error rates were close to that of the maximum-likelihood detection algorithm. The system used T-BLAST and D-BLAST to obtain a lower bit-error rate. Although the bit-error rate was not as low as that of D-BLAST, its coding complexity was much lower.

Vertical Layered Space–Time Coding

To analyze the bit-error performance of different decoding algorithms in atmospheric laser communication after using V-BLAST layered code, the Monte Carlo method was used for system simulation. The effects of the number of receiving and transmitting antennas, noise, and atmospheric flicker on the bit-error rate were analyzed. The performance of the system was compared to that of an uncoded atmospheric (1×1) laser communication system [79–82].

The simulation conditions were as follows:

1. The total transmission power E_S is unchanged and $E_S = 1$ (normalized value).
2. We assume that the channel characteristics and noise can be accurately estimated.
3. The photoelectric-conversion efficiency $\eta = 0.5$.
4. The light-intensity flicker follows a logarithmic normal distribution, and the scintillation index SI = 0.6 and 1.0.
5. The signal is fully synchronized.

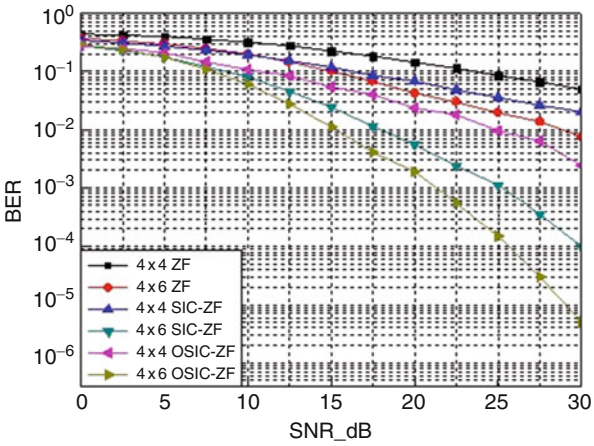


Fig. 20.32 Simulation of ZF, SIC-ZF, and OSIC-ZF

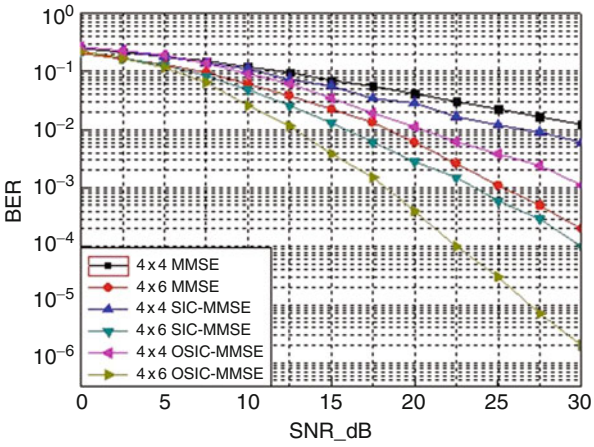


Fig. 20.33 Simulation of MMSE, SIC-MMSE, and OSIC-MMSE

6. The system adopts Q-PPM modulation, where $Q = 4$.
7. The numbers of antennas are 1×1, 1×2, 1×4, 2×2, 2×3, 4×4, and 4×5.

Performance Analysis of the Maximum-Likelihood Decoding Algorithm

Figure 20.34 shows the curves of the bit-error rate and signal-to-noise ratio of the maximum-likelihood decoding algorithm when $SI = 0.6$. The following can be seen from the figure:

1. The values are relative to 1×1. The bit-error performances of the other systems have significantly improved. This shows that layered space–-time coding can effectively overcome the scintillation effect caused by atmospheric turbulence.

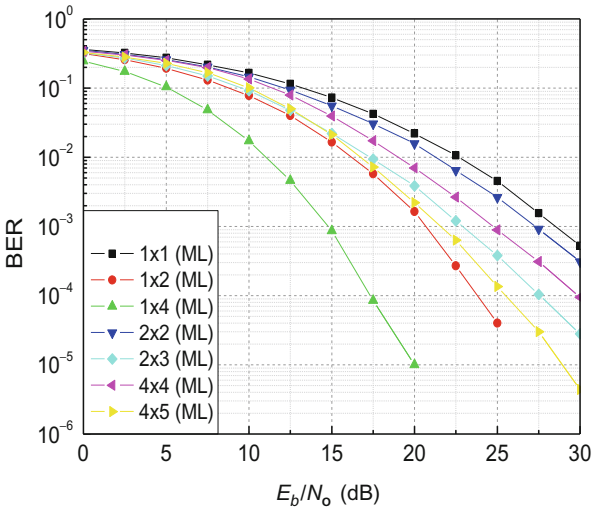


Fig. 20.34 BER and SNR curves of maximum-likelihood decoding algorithm when SI = 0.6

2. The diversity gain is related to the number of detectors. The greater the number of detectors, the greater the diversity gain, and the more obvious the improvement in the bit-error performance. When the bit-error rate is 10^{-3}, relative to the 1×1 system, the signal-to-noise ratios of the 1×2 and 1×4 systems are improved by 6 dB and 14 dB, respectively. When the bit-error rate is 10^{-3}, compared to the 2×3 and 2×2 systems, the signal-to-noise ratio is improved by approximately 4.5 dB.
3. When the number of detectors is the same, the system with more lasers exhibited poorer performance than the system with fewer lasers. This is because vertical layered space–time coding is not based on transmission diversity; thus, it does not obtain transmission-diversity gain, and more importantly, the spatial interference increases.

Figure 20.35 shows the curves of the BER and SNR of the maximum-likelihood decoding algorithm when SI = 1. It can be seen that with the increase in the scintillation index in the 1×1 system, the improvement in the signal-to-noise ratio of the same system is obviously increased. When the bit-error rate is 3×10^{-2}, relative to the 1×1 system, the signal-to-noise ratios of the 1×2 and 1×4 systems improved by approximately 8 dB and 15 dB, respectively. In particular, the 2×2, 2×3, 4×4, 4×5, and other systems have greater diversity gains. This shows that when the scintillation is stronger, the diversity-gain effect of the MIMO system is more evident, resulting in the suppression of the light-intensity fluctuation.

Performance Analysis of Linear-Decoding Algorithms

The linear-decoding algorithms include the zero-forcing and minimum mean-squared error algorithms.

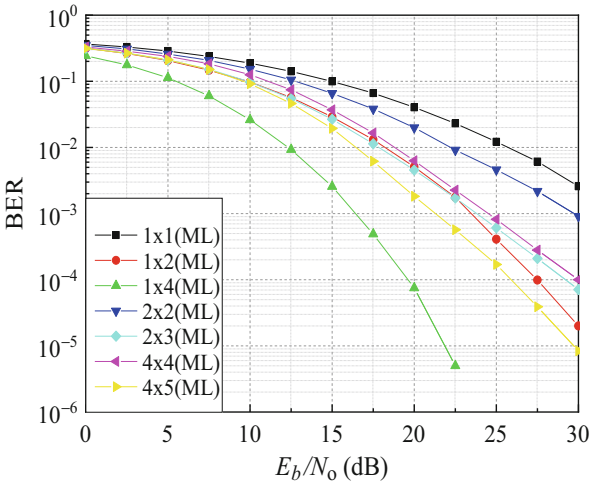


Fig. 20.35 BER and SNR curves of maximum-likelihood decoding algorithm when SI = 1.0

Zero-Forcing Decoding Algorithm

To better describe the bit-error performance of layered space–time coding using the zero-forcing decoding algorithm (ZF), we conducted a simulation analysis. The results are shown in Figs. 20.36 and 20.37. Figure 20.36 shows the error-rate and signal-to-noise ratio curves of the ZF and ML decoding algorithms when SI = 0.6.

It can be seen from the figure that with an increase in the number of lasers and detectors, the bit-error rate of the system decreases gradually with an increase in the signal-to-noise ratio, after using the ZF decoding algorithm. This shows that it is feasible to apply the ZF decoding algorithm to V-BLAST decoding in atmospheric laser communication. However, there are some differences between the rule presented here and the maximum-likelihood decoding algorithm (ML):

1. After using the ZF decoding algorithm, the bit-error performance is obviously improved, owing to the increase in the number of detectors. When the bit-error rate is 6×10^{-2}, relative to the 4×4 and 4×5 systems, the signal-to-noise ratio of the system improves by 7.5 dB. Relative to the 2×2 and 2×3 systems, the signal-to-noise ratio of the system is improved by about 12.5 dB.
2. This is because an increase in the number of detectors is equivalent to an increase in the receiving aperture, indicating that the system can effectively overcome the scintillation effect caused by atmospheric turbulence.
3. After using the ZF decoding algorithm, the BERs of the 4×4 and 4×5 systems are worse than those of the 2×2 and 2×3 systems. This is due to the increase in spatial interference between signals after multiplexing technology is adopted at the transmitter.

Figure 20.37 shows the BER and SNR curves of the ZF and ML decoding algorithms when SI = 1.0. It can be observed from the figure that the law reflected

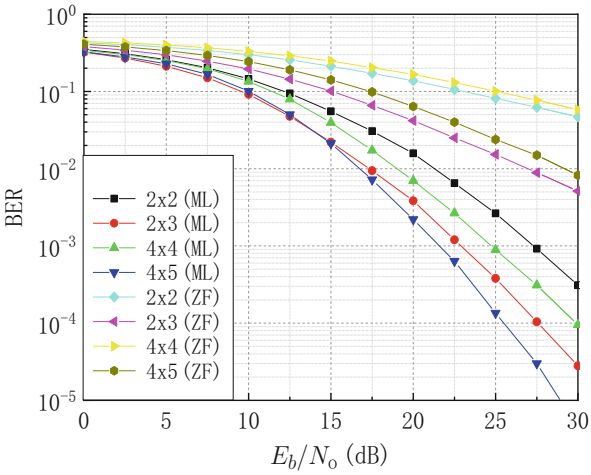


Fig. 20.36 BER curves of ZF and ML decoding algorithms at SI = 0.6

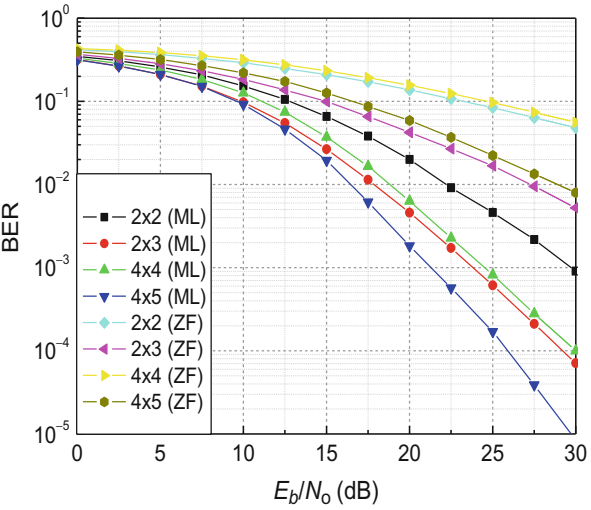


Fig. 20.37 BER curves of ZF and ML decoding algorithms when SI = 1.0

at this time is the same as that in Fig. 20.36; however, the differences are only 2 × 2 and 4 × 4, and 2 × 3 and 4 × 5, respectively. The performance gap between these systems was significantly reduced. This again shows that when the flicker is stronger, the diversity-gain effect of the MIMO system is more evident.

Minimum Mean-Squared Error Decoding Algorithm

Figure 20.38 shows the error-rate curves of the MMSE and ML decoding algorithms when the scintillation index is 0.6. It can be seen from the figure that with an increase in the number of lasers and detectors, the bit-error rate of the system decreases

gradually with an increase in the signal-to-noise ratio after using the MMSE decoding algorithm.

Figure 20.39 shows the BER and SNR curves of the MMSE and ML decoding algorithms when SI = 1.0. The rule reflected in this figure is very similar to that in Fig. 20.53; however, the improvement in the system bit-error rate is different. The improvement in the system bit-error rate is especially greater in the case of a strong scintillation, compared to a classic single-input single-output (SISO) system, when the minimum mean-squared error decoding algorithm is adopted at 2 × 2 and 4 × 4.

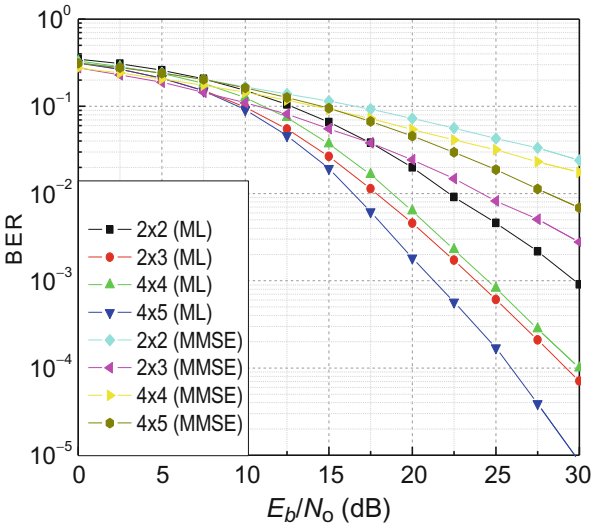


Fig. 20.38 BER and SNR curves of MMSE and ML decoding algorithms when SI = 1.0

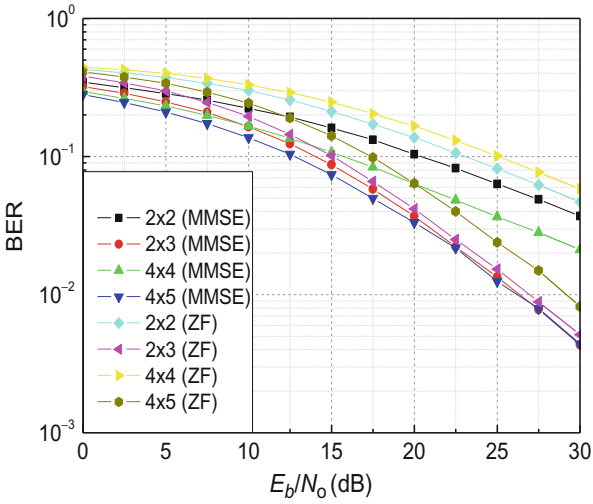


Fig. 20.39 Comparison of ZF and MMSE decoding algorithms when SI = 0.6

The BER characteristics of the different decoding algorithms under different turbulence conditions were compared, and the decoding complexity was analyzed. The results show that the performance of the diagonal layered space–time code is the best of the three codes, followed by the spiral-layered code and horizontal layered space–time code. The bit-error rate of the diagonal layered space–time code does not change significantly when the turbulence is enhanced. The largest change is in the horizontal layered space–time code, which also shows that the diagonal layered space–time coding system can suppress the influence of strong atmospheric turbulence on the system.

Whether in weak or strong turbulence, when the number of transmitting and receiving antennas increases gradually, the best performance of the three-layered space–time codes is the diagonal layered space–time code, followed by the spiral layered space–time code, and finally, the horizontal layered space–time code. The horizontal layered space–time code is not coded in the spatial domain; therefore, its performance is poor. Compared with diagonal layered space–time coding, spiral layered space–time coding has the advantages of lower coding complexity and no redundancy, which makes better use of space and time.

Although the performance of diagonal layered space–time coding is optimal and makes full use of the resources in the space and time domains, because it has $n\ (n-1)/2$ redundancy, the coding increases the redundancy and obtains higher coding performance. Cascaded layered space–time coding is also a direction for space–time code development.

Space–Time Codes

Space–time block codes and space–time trellis codes are based on transmit diversity. To further improve the performance of space–time codes and suppress atmospheric-turbulence effects, we studied differential space–time codes. In this section, we introduce our work on orthogonal space–time codes, space–time trellis codes, hybrid space–time codes, differential space–time codes, and unitary space–time codes [80, 83–87].

Orthogonal Space–Time Coding

The earliest space–time coding (STC) structure was proposed by Alamouti [88], and it has attracted extensive attention. The original form of the space–time block code [89] was also proposed by Alamouti. Trokh extended the two-antenna transmission-diversity scheme proposed by Alamouti [90] to any number of transmit antennas to obtain a full-diversity gain and simple decoding. Simon and Vilnrotter et al. improved the classic Alamouti code [89] and proposed an orthogonal space–time coding scheme suitable for IM/DD optical communication.

Traditional Alamouti codes include negative and complex signals; however, complex conjugations and negative numbers cannot be represented by “on” and “off.” Because optical wireless communication generally adopts intensity modulation, the

traditional Alamouti code [89] has been improved and a space–time coding method suitable for an IM/DD atmospheric turbulence channel has been proposed.

We combined space–time coding with modulation technology, and used the relationship between binary bits (i.e., the complementary relationship between binary bits "0" and "1") to construct the bit-complement orthogonal space–time block code (BCOSTBC), which first performs a bit complement before Q-PPM modulation. When the total power E_s of the system is constant and the channel parameters are known, the Monte Carlo method is used to simulate the bit-error performance of 1×1, 1×2, 2×1, and 2×2. The simulation parameters are $\eta = 0.5$ and $R = 1$. The simulation results are shown in Figs. 20.40, 20.41, 20.42, and 20.43.

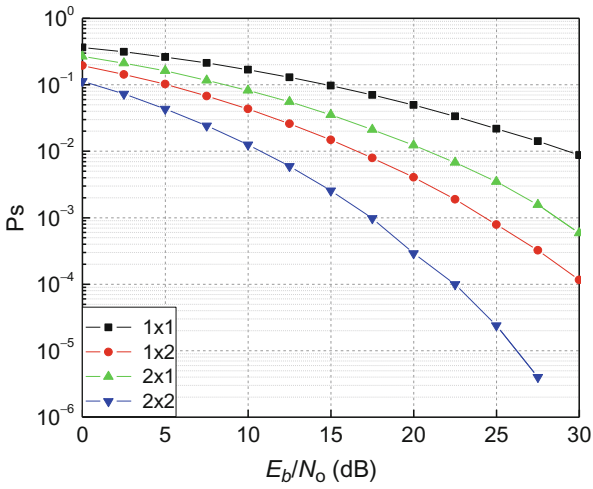


Fig. 20.40 SI = 0.6, word-error rate of BCOSTBC code with 4-PPM

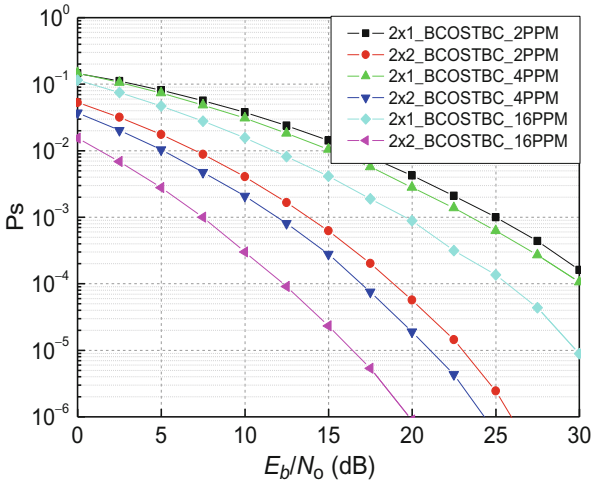


Fig. 20.41 SI = 1.0, word-error rate of BCOSTBC code with different modulation modes

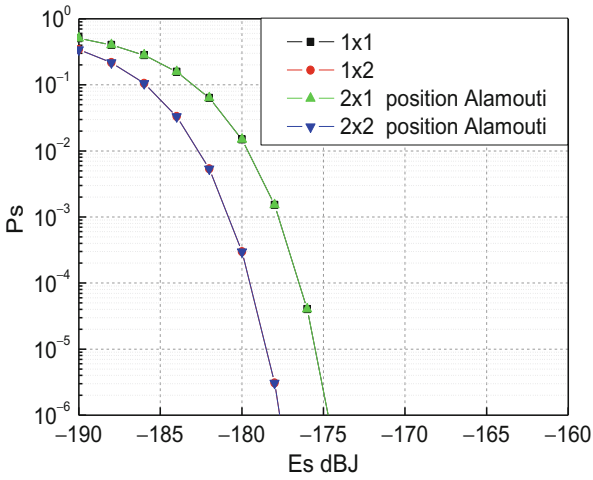


Fig. 20.42 Bit-error rate without background radiation and fading

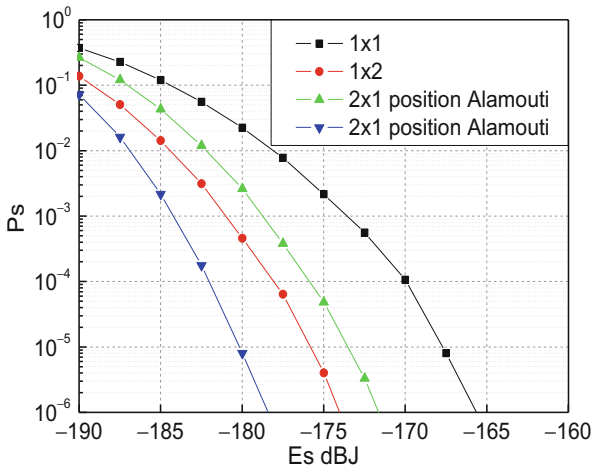


Fig. 20.43 Bit-error rate with fading but without background radiation

Figure 20.40 shows the bit-error rate of the BCOSTBC code modulated by 4-PPM when SI = 0.6. As shown in Fig. 20.40, the error-code performance of the system is significantly better after BCOSTBC coding. When $P_s = 10^{-2}$, relative to 1×1, the BCOSTBC code is adopted after 2×1 and 2×2. Then, the bit-error performance of the system is improved by approximately 8 dB and 18 dB, respectively.

Figure 20.41 shows the error rate of the BCOSTBC code for different modulation modes when SI = 0.6. As shown in Fig. 20.41, under different modulation modes, the error rate of the system decreases with an increase in the signal-to-noise ratio. Moreover, in the same system, the word-error rate of the system decreases with an increase in Q.

Different pulse positions were used to represent the complex conjugate and negative symbols. For example, when $Q = 4$, x [1 0 0 0], x * [0 0 0 1], −x * [0 0 1 0], and −x [0 1 0 0]; the encoding matrix is converted into a real matrix represented by the pulse position. Thus, it can be transmitted using IM/DD-type optical communication.

The Monte Carlo method was used to simulate the bit-error performance of 1 × 1, 1 × 2, 2 × 1, and 2 × 2. While the total power E_s of the system remains unchanged, the simulation parameters are SI = 0.6, wavelength $\lambda = 1550$ nm, $\eta = 0.5$, and $Q = 4$.

Figure 20.42 shows the bit-error performance without background radiation or fading. The following can be seen from Fig. 20.42:

1. The system's bit-error performance in 2 × 2 and 1 × 2 is obviously better than that in 1 × 1. When $P_s = 10^{-6}$, relative to 1 × 1, the bit-error performance of the 2 × 2 and 1 × 2 systems was approximately improved. This shows that increasing the number of detectors is equivalent to increasing the receiving aperture, which is conducive to improving the system performance.
2. The error-rate curves of the 1 × 1 and 2 × 1 systems coincide. This indicates that the bit-error performance of the system does not improve with an increase in the number of lasers when the total power remains unchanged.

Figure 20.43 shows the bit-error performance when there is no background radiation or fading, assuming the scintillation index SI = 0.6. The following can be seen from Fig. 20.43:

1. When the number of lasers remains unchanged, increasing the number of detectors is conducive to improving the bit-error performance of the system.
2. When the number of lasers increases, the bit-error performance of the system does not improve. This is because, when the total power is constant, increasing the number of lasers M means that the power of each laser will be reduced to $1/M$, and the background radiation introduced with each character will be M times that of the system.

Space–Time Trellis Coding

In 1998, Tarokh, Seshadri, and Calderbank [91] proposed the space–time trellis code (STTC) after jointly designing error-control coding, modulation, transmit, and receive diversity technologies. The STTC design criteria were based on flat slow-fading and fast-fading channel assumptions.

We analyzed the change of the bit-error rate with the SNR when two transmitting antennas were used with different numbers of receiving antennas. The simulation conditions were as follows:

1. The channel characteristics and noise at the receiving end can be accurately estimated.

2. The photoelectric-conversion efficiency is set to 0.8, and the scintillation index SI is set to 0.7 and 0.9. (The SI generally ranges from 0.4 to 1.)
3. We assume that the transmitted signals of each antenna arrive at the receiving antenna synchronously and simultaneously.
4. Four-way PPM was adopted for system modulation.
5. When analyzing the bit-error rate, the antenna numbers are 2×2, 2×3, and 2×4.
6. The coding coefficient $g1 = [0\ 2;\ 2\ 0]$; $g2 = [0\ 1;\ 1\ 0]$.

The BER curves for different receiving antennas and scintillation indexes are shown in Figs. 20.44 and 20.45. It can be seen from Figs. 20.45 and 20.46 that with the increase in the number of receiving antennas, the bit-error rate performance is

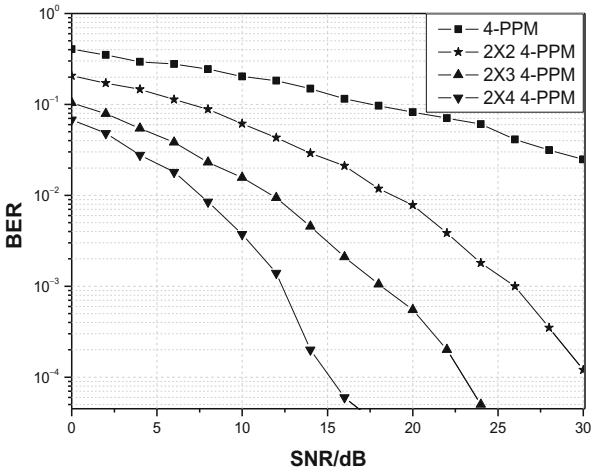


Fig. 20.44 BER and SNR curves when SI = 0.7

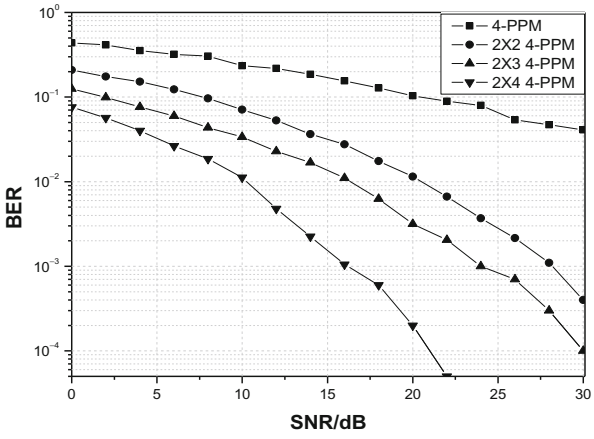


Fig. 20.45 BER and SNR curves when SI = 0.9

significantly improved. Moreover, the STTC 4-PPM system achieves a better BER performance than 4-PPM modulation under the same noise condition. Comparing Figs. 20.45 and 20.47, it can be seen that when SI = 0.9 and SNR = 20 dB with a system ratio of 2 × 3, the bit-error rate of the system is an order of magnitude higher.

In a optical wireless STTC system with three lasers, the influence of the number of receiving antennas under three transmitting antennas on the bit-error rate with the change in the SNR is analyzed. The simulation conditions are as follows:

1. The photoelectric conversion efficiency is set to 0.8, and the scintillation index SI is set to 0.7.
2. We assume that the transmitted signals from each antenna arrive at the receiving antenna synchronously and simultaneously.
3. Four-way PPM is adopted for system modulation.

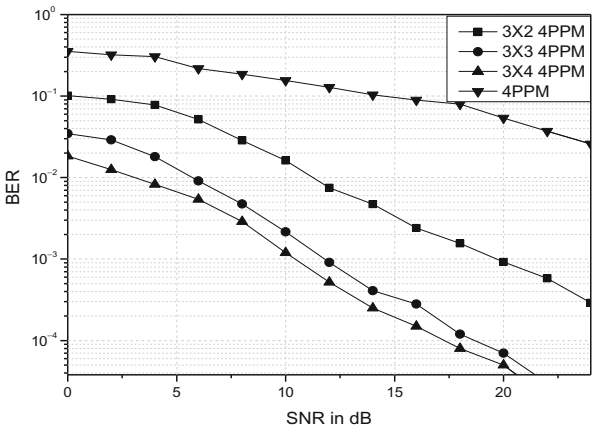


Fig. 20.46 BER and SNR curves when SI = 0.7

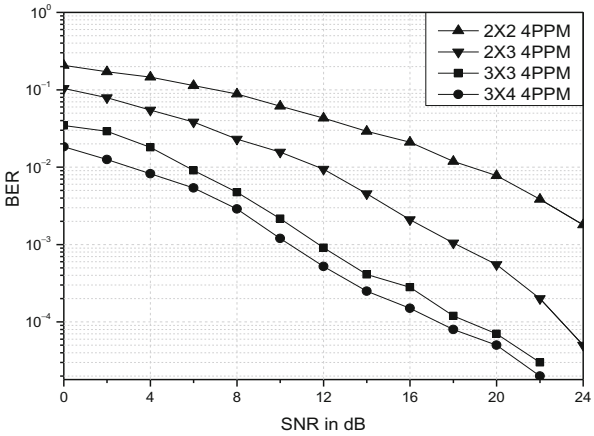


Fig. 20.47 Comparison of BER between a three-antenna system and two-antenna system when SI = 0.7

4. When analyzing the bit-error rate, the number of antennas is 3×2, 3×3, and 3×4.
5. $g1 = [0\ 2\ 2;\ 1\ 2\ 3]$ and $g2 = [2\ 3\ 3;\ 2\ 0\ 2]$.

As shown in Fig. 20.45, the three-antenna system has the same advantages as the two-antenna system. In Fig. 20.46, a direct comparison is made between the two transmitting-antenna system and the three transmitting-antenna system. It can be seen that when the number of receiving antennas is three and SNR = 20 dB, the 3×3 system ratio is improved by two orders of magnitude. The BER decreases with an increase in the number of transmitting antennas.

Differential Space–Time Coding

In 2000, Tarokh and Jafarkhani proposed a two-antenna differential space–time block code based on an orthogonal design [92], which was an improvement on the Alamouti orthogonal space–time block code [93]. Compared to the traditional two antenna space–time block code (Alamouti code), this scheme has a 3-dB performance loss [93] and quadrature phase shift keying (QPSK) modulation has a problem with signal expansion. Tarokh and Jafarkhani generalized the results in [94] such that they could be applied to multiple transmission antennas [94].

In 2000, Hochwald and Hughes successively proposed differential unitary space–time modulation [95] and differential space–time modulation [96]. Differential unitary space–time modulation introduced a type of diagonal matrix, and the symbols on the diagonal of the matrix are phase-shift keying (PSK)-modulated signals. In addition, they proposed a differential space–time modulation scheme for multiple transmit antennas, based on group codes [96].

We define the received energy of each symbol of each detector as $E_S = (\eta I_S)^2 T_S$, when there is no turbulence, where $T_S = T/Q$ is the time-slot length, Q is the modulation order, and T is the character period. It is assumed that the total power of the system is normalized to $E_s = 1.0$.

According to Taylor's frozen turbulent-atmosphere theory, the channel coefficients are assumed to remain unchanged in two adjacent transmission-code matrices. The channel-fading coefficient remains unchanged in two consecutive transmission-matrix blocks; that is, the channel is a slow-fading channel. The BER of the system was simulated using the Monte Carlo method: 1×1, 2×2, 2×3, 4×1, 4×2, 4×3, and 4×4. $\eta = 0.5$. A binary pulse-position modulation (BPPM) mode is adopted, i.e., $Q = 2$, and the binary data stream has one million sample values.

Figure 20.48 shows the bit-error rate of differential space–time coding of two transmitting antennas when the scintillation index SI is 0.4, using the lognormal distribution model. It can be observed from the figure that the system performance gradually improves with an increase in the number of receiving apertures. When the bit-error rate is 10^{-3}, the two receiving apertures are reduced by about 8.5 dB compared with a single one. Compared with one and two receiving apertures, the

bit-error rates of three receiving apertures are reduced by about 13 dB and 4.5 dB, respectively.

Figure 20.49 shows the bit-error rate of the differential space–time coding of two transmitting antennas when the turbulence model adopts the gamma–gamma distribution model and a scintillation index SI = 0.9. It can be seen that when the bit-error rate is 10^{-2}, the two receiving apertures are reduced by about 10 dB compared with a single one. Moreover, the three receiving apertures are reduced by about 15 dB and 5 dB compared with one and two receiving apertures, respectively.

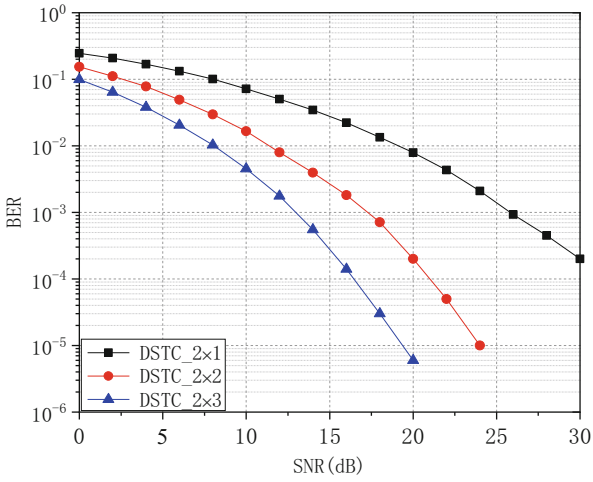


Fig. 20.48 Bit-error rate of differential space–time coding with SI = 0.4 and BPPM modulation

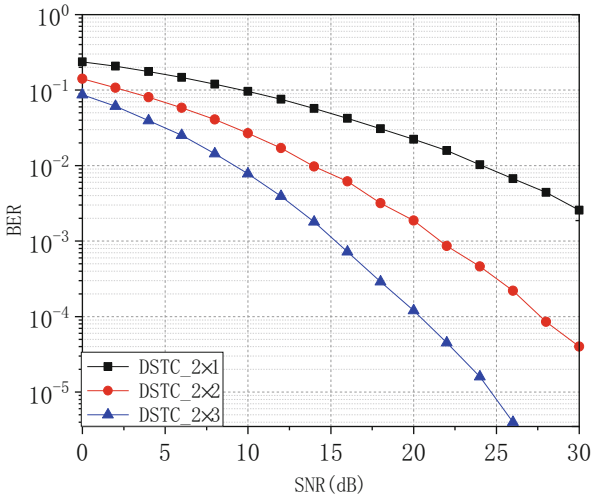


Fig. 20.49 Bit-error rate of differential space–time coding with BPPM modulation at SI = 0.9

Figure 20.50 shows the bit-error rate of differential space–time coding with two transmitting antennas when the scintillation index SI = 1.12 and a gamma–gamma-distribution turbulence model is adopted. When the bit-error rate is close to 10^{-2}, the two receiving apertures are reduced by about 14 dB, compared with a single one. Compared with one and two receiving apertures, three receiving apertures are reduced by about 20 dB and 6 dB, respectively.

Figure 20.51 compares the bit-error rate of differential space–time code with the bit-error rate of Alamouti code [97] improved by Simon and Vilnrotter. A gamma–gamma distribution model is adopted and the scintillation index SI = 0.9.

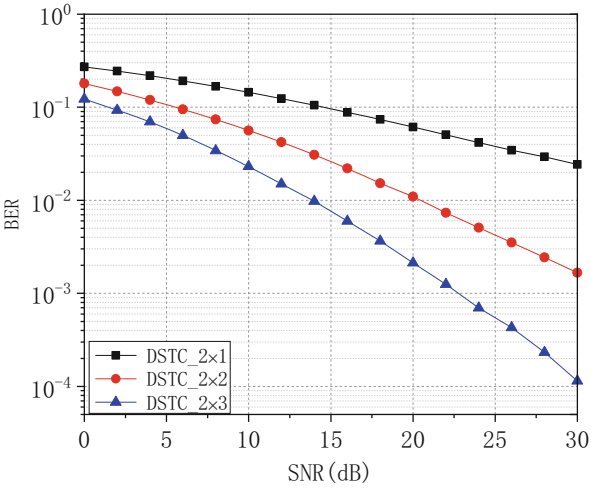


Fig. 20.50 Bit-error rate of differential space–time coding with BPPM modulation at SI = 1.12

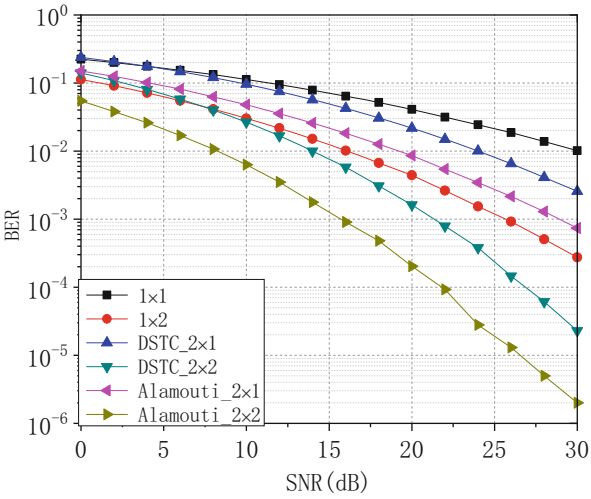


Fig. 20.51 Comparison of bit-error rates of differential space–time code and Alamouti code with BPPM modulation at SI = 0.9

It can be seen from the figure that the performance curves of the differential space–time coding method and the improved Alamouti coding method are parallel, which shows that the differential scheme can achieve complete transmission diversity with the improved Alamouti coding [97]. When the bit-error rate is 10^{-2}, the difference between the differential coding and the improved Alamouti coding [97] is about 5 dB; however, the differential coding scheme in the 2×1 and 2×2 systems is significantly better than the one without coding in the 1×1 and 1×2 systems.

Figure 20.52 compares the bit-error rate of the differential space–time code between two and four transmitting antennas when the scintillation index SI is 0.4 under a lognormal-distribution turbulence model. We can see the following from the figure:

1. When the bit-error rate is 10^{-3}, the bit-error performance of the 4×1 system is improved by about 6 dB, relative to the 2×1 system. When the bit-error rate is 10^{-4}, the bit-error performance of the 4×2 system is improved by about 3.5 dB, relative to that of the 2×2 system.
2. When the bit-error rate is 10^{-3}, the 2×2 system's bit-error performance is improved by about 8 dB over the 4×1 system. When the bit-error rate is 10^{-4}, the bit-error performance of the 4×2 system is improved by approximately 7.5 dB, relative to the 4×1.
3. For two different MIMO systems, when the product MN of the number of transmitting antennas M and the number of receiving antennas N is equal, the bit-error performance of the differential space–time code of the 2×2 system is obviously better than that of the 4×1 system. When the bit-error rate is 10^{-4}, the bit-error performance of the 2×2 system is improved by approximately 3 dB relative to the 4×1 system.

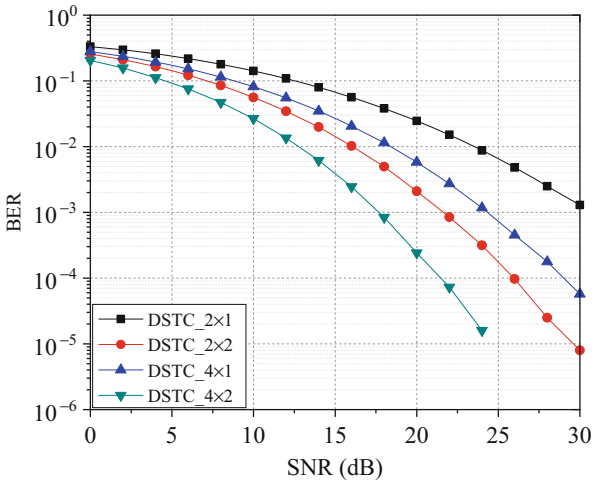


Fig. 20.52 Comparison of bit-error rates of differential space–time codes with two and four transmitting antennas when SI = 0.4

Figure 20.53 shows the bit-error rate for two and four antennas when the scintillation index SI = 0.9 using the gamma–gamma distributed turbulence model. The performance shown in Fig. 20.52 is significantly lower than that shown in Fig. 20.51, owing to the influence of the atmospheric-turbulence intensity.

The performance shown in Fig. 20.54 was significantly worse than that shown in Fig. 20.55 and Fig. 20.53. It can be seen that when the differential space–time code adopts a 4 × 2 system relative to the 2 × 2, 4 × 1, and 2 × 1, the gap increases significantly. When the signal-to-noise ratio is 20 dB, the bit-error rate performance of the 4 × 2 system is improved by one order of magnitude, relative to the 2 × 2

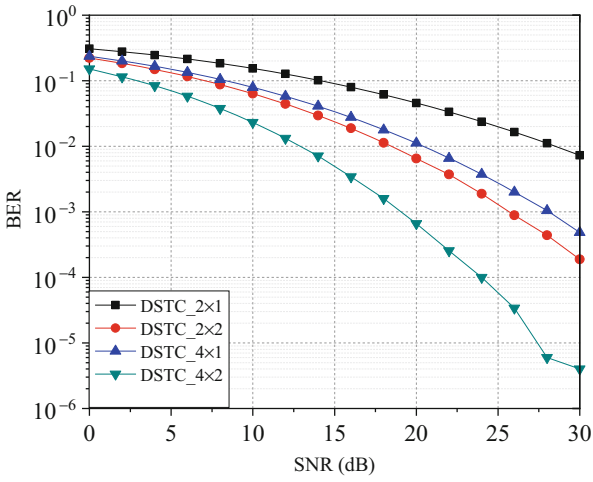


Fig. 20.53 Comparison of bit-error rates of differential space–time codes with two and four transmitting antennas when SI = 0.9

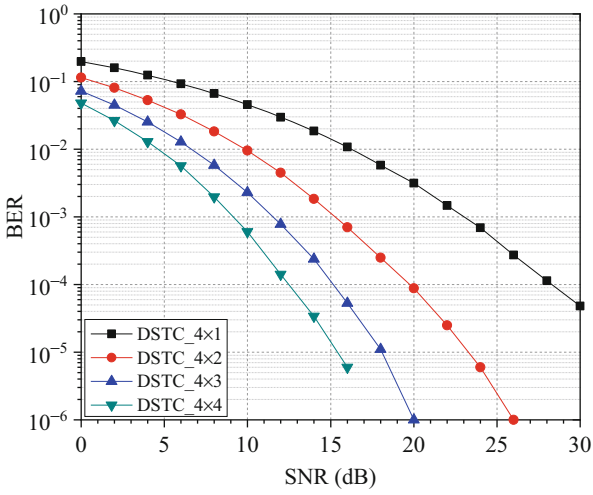


Fig. 20.54 Bit-error rate diagram of differential space–time codes with four transmitting antennas and different receiving antennas

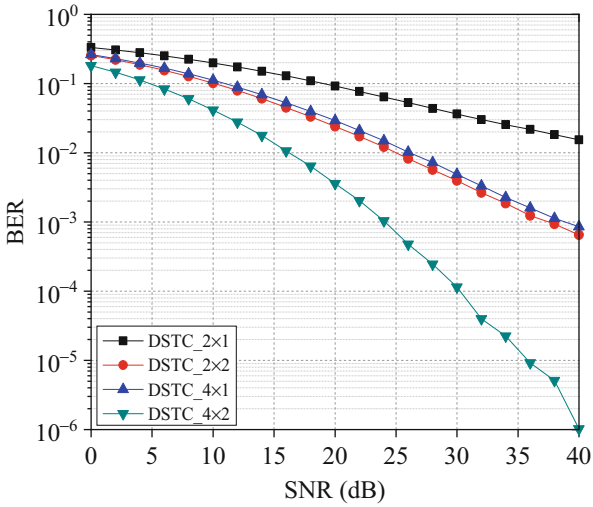


Fig. 20.55 Comparison of bit-error rates of space–time codes with two and four transmitting antennas

system. When the signal-to-noise ratio is 40 dB, the 4 × 2 system is improved, relative to the 2 × 2 system, and the bit-error rate performance of the system is improved by three orders of magnitude.

Figure 20.55 shows the curve of the bit-error rate versus the signal-to-noise ratio of the differential space–time code when the number of transmitting antennas is four and the number of receiving antennas is one, two, three, and four, when a lognormal-distribution turbulence model is adopted and the scintillation index SI = 0.6. It can be seen from the figure that with the increase in the number of receiving antennas, the system performance is better.

When the signal-to-noise ratio is 15 dB, the bit-error rate performance of the 4 × 4 system improved by three orders of magnitude relative to the 4 × 1 system. The bit-error rate performance of the 4 × 4 system improved two orders of magnitude relative to the 4 × 2 system. The bit-error rate performance of the 4 × 4 system improved by one order of magnitude relative to the 4 × 3 system.

Unitary Space–Time Coding

It is difficult, and sometimes even unrealistic, to effectively estimate a channel in a mobile environment with rapidly changing channel characteristics. Therefore, the study of space–time coding without channel estimation at the receiver has great practical value. Hughes, Hochwald, Marzetta, and other scholars [98–100] studied blind space–time block codes without channel-parameter estimation at the receiver.

Hochwald et al. [96] deduced the unitary space–time code (USTC) according to the channel capacity of a Rayleigh fading channel in information theory. If the sending code block is a unitary matrix, the receiver can perform maximum-likelihood decoding without knowing the channel state.

We assume that the total transmission power E_s is unchanged, and the normalized value of E_s is considered. According to Taylor's frozen-field theory, it is assumed that the channel parameters remain unchanged in two consecutive transmission-code matrix blocks. The photoelectric conversion efficiency $\eta = 0.5$ and the number of samples is 100,000. When the scintillation index SI of the channel was 0.4 (weak turbulence) and 0.8 (medium turbulence), the above two methods were used to test the bit-error rate of the 4×2 and 4×1 systems. The simulation results are shown in Figs. 20.56, 20.57, 20.58, and 20.59.

As shown in Fig. 20.58, when SI = 0.4 (weak turbulence) and the bit-error rate is 10^{-3}, the signal-to-noise ratio difference of the 4×2 and 4×1 systems is 5 dB. This shows that the performance of the unitary space–time coding scheme is greatly improved with an increase in the receiver-aperture number.

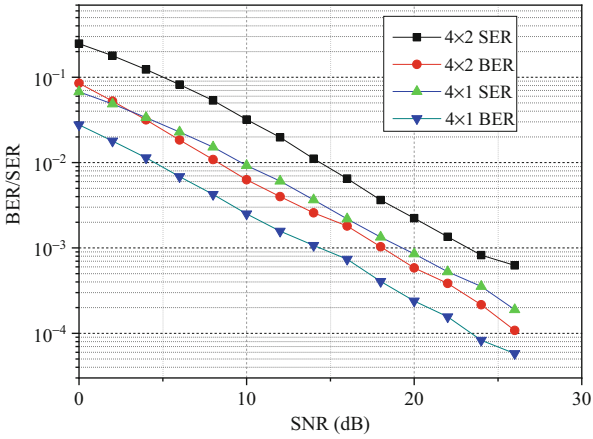


Fig. 20.56 BER performance of unitary space–time coding with channel estimation at SI = 0.4

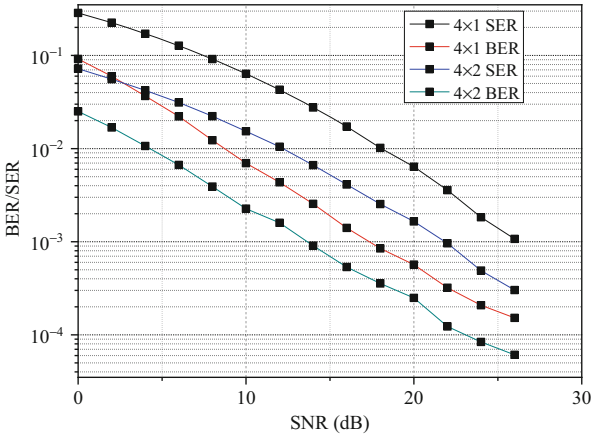


Fig. 20.57 BER performance of unitary space–time coding with channel estimation at SI = 0.8

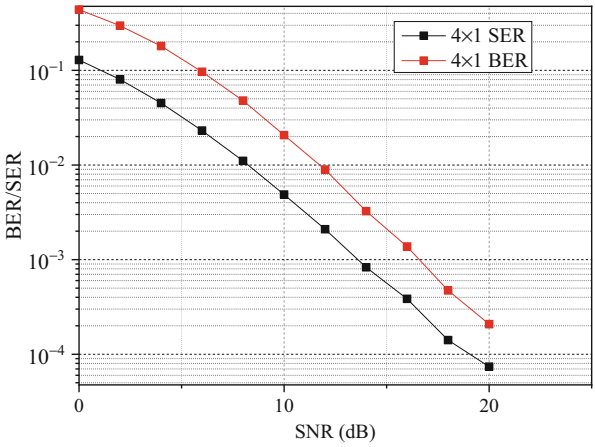


Fig. 20.58 BER performance of unitary space–time coding without channel estimation at SI = 0.4

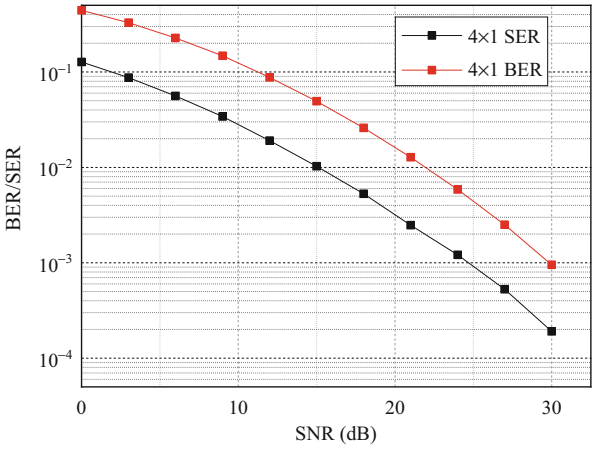


Fig. 20.59 BER performance of unitary space–time coding without channel estimation at SI = 0.8

As shown in Figs. 20.58 and 20.57, when SI = 0.4 and the bit-error rate is 10^{-3}, the 4 × 1 system signal-to-noise ratio remains almost constant at 19 dB, and the 4 × 2 system signal-to-noise ratio is 15 dB. When SI = 0.8 (medium turbulence) and the bit-error rate is 10^{-3}, the 4 × 1 system SNR is 23 dB, which is 4 dB different from that of SI = 0.4. The 4 × 2 system SNR is 14 dB, which is 1 dB different from that of SI = 0.4.

Comparing Figs. 20.58 and 20.59, we can see that when the signal-to-noise ratio is 25 dB, in weak or medium turbulence (SI = 0.4 or 0.8), the 4 × 1 system bit-error rate is 10^{-3} to 10^{-4}, the bit-error rate of the 4 × 2 system is lower, and the bit-error rates of the system under the two turbulence conditions are similar.

By comparing the system-simulation diagrams in Figs. 20.50 and 20.51, which do not require channel estimation, it can be seen that when the signal-to-noise ratio is

20 dB and SI = 0.4 in weak turbulence, the bit-error rate of the 4 × 1 system is low, and when SI = 0.8, the bit-error rate of the 4 × 1 system is between 10^{-3} and 10^{-4}, which is obviously inferior to the bit-error performance that requires channel estimation. This indicates that a system requiring channel estimation has a strong ability to suppress the influence of turbulence.

Hybrid Space–Time Coding

Layered space–time coding (LSTC) [101–103] and space–time block coding (STBC) [104–106] are two important space–time coding methods. Although LSTC has poor bit-error performance, it can achieve high spectrum utilization. Although the STBC code uses orthogonality to achieve transmission diversity in the system, its performance is significantly better than that of the LSTC. However, because of redundant information in the coding matrix, the transmission efficiency is low [107, 108]. Therefore, improving the transmission rate of the STBC code is the first problem to be solved. We make a reasonable compromise between multiplexing gain and diversity gain and propose a hybrid space–time coding scheme suitable for IM/DD optical communication [109].

When the total transmission power remained unchanged and the channel-attenuation coefficient was determined, the Monte Carlo method was used to simulate the bit-error performance of a 4 × 4 system. The simulation parameters are $Q = 4$, SI = 0.6, $\eta = 0.5$, and a 2 × 2 mixed coding scheme is used. The simulation results are shown in Figs. 20.60 and 20.61.

Figure 20.60 shows the bit-error performances of different systems using different coding methods. The following can be seen from Fig. 20.59:

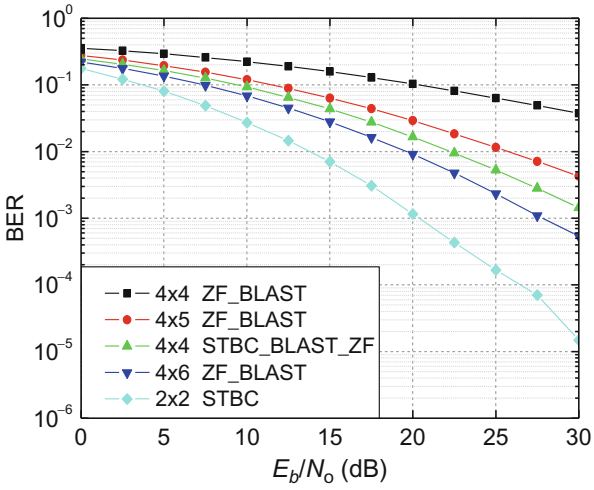


Fig. 20.60 BER performance of different systems with different coding methods

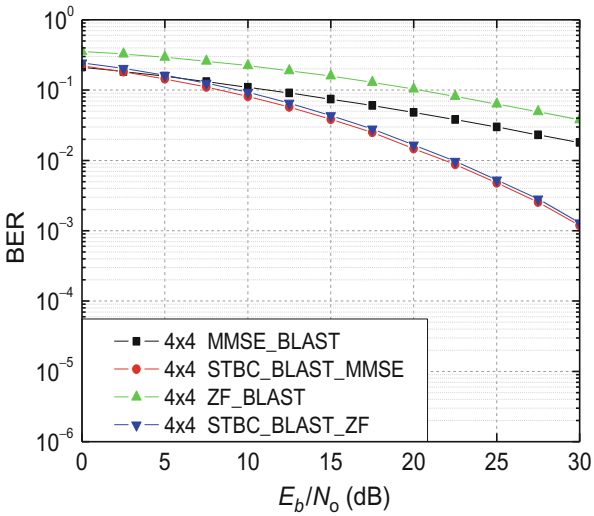


Fig. 20.61 BER performance of two systems using different decoding algorithms

1. When comparing a 2 × 2 STBC system and a 4 × 4 ZF_ BLAST system, it can be seen that there is a wide range between the error rates, and the error-rate curve of the STBC system is steeper than that of the BLAST system. This indicates that the STBC system can obtain a larger diversity gain without a partial multiplexing gain.
2. Comparing the bit-error performance of the hybrid space–time coding system between a 2 × 2 STBC system and a 4 × 4 ZF_ BLAST system, it can be seen that the performance of the hybrid space–time coding is significantly better than that of traditional BLAST codes. When the bit-error rate is 10^{-2}, relative to the 4 × 4 ZF_ BLAST system, the performance of the 4 × 4 hybrid space–time coding system is improved by approximately 14 dB. This shows that introducing the STBC code into layered space–time coding increases the diversity of the first translated layer, ensures the correctness of the first layer, and reduces error propagation.
3. The error performance of 4 × 4 hybrid space–time coding system is between that of a 4 × 5 ZF_ BLAST and that of a 4 × 6 ZF_ BLAST system. We note that the 4 × 4 hybrid space–time coding system obtains both reception-diversity gain and transmission-diversity gain.

The MMSE and ZF detection algorithms were used to analyze the traditional BLAST system and hybrid space–time coding system, and their bit-error performance is shown in Fig. 20.61. The following can be seen from the figure:

1. When the MMSE decoding algorithm is used in the hybrid space–time coding system, the bit-error performance of this scheme is obviously better than that of the traditional BLAST system.

2. In a traditional BLAST system, the performance of the MMSE algorithm is better than that of the ZF algorithm because it considers the influence of noise on the signal.

   However, in the hybrid space–time coding system, the performances of the two algorithms were very similar. When the bit-error rate was 2×10^{-3}, the bit-error performances of the two algorithms were the same.

Summary

In 2011, Qiu et al. [110] built a fiber-channel MIMO model using the modal dispersion of a multimode fiber. In 2011, Li et al. [111] performed a theoretical analysis of vertical layered space–time coding (V-BLAST) using the zero-forcing (ZF) detection, minimum mean-squared error (MMSE) detection algorithm, and maximum-likelihood (ML) detection algorithm for fiber-channel MIMO. The results showed that the ZF algorithm performed poorly in multimode fiber links and the MMSE algorithm performed better than the ZF algorithm; the ML algorithm obtained the minimum error probability of the original symbol, and its performance was the best. In 2011, Zhao et al. [112] analyzed the channel capacity of optical wireless space–time coding, and Xing [113] studied quasi-orthogonal space–time coding.

Space–time coding has gradually attracted attention in the field of optical wireless communication. Since 2007, the author has devoted himself to research on space–time coding in optical wireless communication [8–85]. Drawing on previous research results of space–time coding in the radio-frequency field, the author extended it to the field of optical wireless communication. Future directions include adaptive coding according to the channel changes to improve the channel capacity. Error-correction coding is added to adapt to the change in channel characteristics to improve the reliability of optical wireless communication.

Concatenated Code

Studies on space–time concatenated codes mainly investigate concatenation schemes and decoding methods between different codes to reduce the complexity of multistage iterative decoding algorithms [114–118]. Generally, error-correction code is used as the outer code, and space–time code is used as the inner code to construct a concatenated space–time code [119].

In 1966, Forney proposed a serial concatenated-coding technology [120] that used short component codes to construct long codes. It used non-binary and long codes as outer codes, and binary and short codes as inner codes. It achieved better error-code characteristics and decoding complexity within the allowable range.

Since the emergence of wireless laser communication technology, researchers have attempted to improve the performance of communication systems using concatenated-coding technology. In 2012, Xu et al. [121] studied a scheme for

concatenating turbo-product codes and space–time block codes, and the bit-error rate of the concatenated codes was clearly better than that of the space–time codes.

Cascaded Orthogonal Space–Time Coding

Cascaded Orthogonal Space–Time Block-Code Model

Concatenated coding technology uses two or more codes to concatenate and increase the total code length to improve the system performance without increasing the complexity of the system hardware. This method is particularly suitable for improving the performance of atmospheric laser communication systems. Because a non-binary RS code is a maximum-distance separable code that is easy to implement, it is the most natural choice as an external code. Our outer code uses the Reed–Solomon (RS) code, and the inner code uses orthogonal space–time coding. The principle model of cascaded space–time block coding is shown in Fig. 20.62.

After a serial-parallel transformation of the information sequence sent by the source, the sequence first enters the external-code encoder (RS code) for coding, and the external-code encoder outputs non-binary RS-code symbols. Because the PPM modulation symbols match the RS code, the RS code does not need to be converted into binary code, and the output of the external coder can be sent directly to the interleaver for interleaving. The interleaved information symbols are sent directly to the internal-code encoder (STBC encoder) for coding. For orthogonal space–time coding using PPM modulation, the orthogonality is reflected in the pulse position, so it can be directly transmitted by the laser.

The receiver sends the received PPM-modulated symbols to the internal-code decoder for decoding (STBC decoder) and then directly enters the external-code decoder (RS decoder) for decoding after deinterleaving. The original information can be recovered after a parallel-to-serial conversion. In this system, because of the interleaving technology, burst errors can be converted into random errors; thus, the number of scattered error codes falls within the range of error-correction decoding, which further improves the error-correction capability of the system.

Performance Analysis of Cascaded Orthogonal Space–Time Coding

The RS (n, k) code is a type of Bose–Chaudhuri–Hocquenghem (BCH) code with strong error-correction ability. For RS codes that can correct t sign errors, the minimum code distance is $d = 2\,t + 1$, and the number of supervision elements is

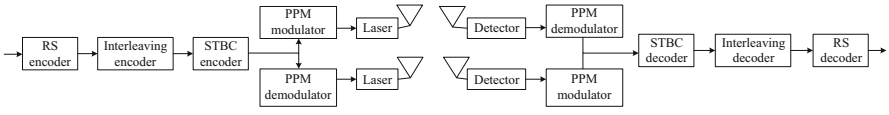


Fig. 20.62 Principle of cascaded space–time block coding

$r = 2\,t$. If the system takes the RS code as the outer code and adopts the errors-only decoding (EOD) algorithm for decoding, the internal-code decoder only needs to output the symbol decision. It does not need to provide symbol deletion or information about the symbol reliability [122].

Thus, if the bit-error rate output by the internal-code decoder is P_S, for an RS (n, k) code that can correct t symbol errors, the upper limit of the bit-error rate output after decoding is [123, 124]

$$P_b < \frac{2^{k-1}}{2^k - 1} \cdot \sum_{j=t+1}^{n} \frac{j+t}{n} \cdot C_n^j \cdot P_{\text{s_BCPSTBC}}^j \cdot \left(1 - P_{\text{s_BCOSTBC}}\right)^{n-j}. \tag{20.1}$$

In this manner, the upper limit of the output bit-error rate P_b of the concatenated code can be estimated from the output bit-error rate $P_{\text{S_BCOSTBC}}$ of the specific internal code. The upper bound of the decoding error rate of the concatenated code can also be estimated according to the relationship between the error rate and the word-error rate:

$$P_s = \frac{2(Q-1)}{Q} P_b. \tag{20.2}$$

With the total power E_s of the system unchanged and the channel parameters known, the Monte Carlo method was used to simulate the bit-error performance of the 1×1, 2×1, and 2×2 systems. Simulation parameter $\eta = 0.5$. The simulation results are shown in Figs. 20.63, 20.64, 20.65, and 20.66.

Figure 20.63 shows the word-error rate of the BCOSTBC and repeat codes [125, 126] after concatenating the RS (15, 7) code when SI = 0.4. To facilitate matching between symbols, 16-PPM modulation was adopted. The following can be observed from the figure:

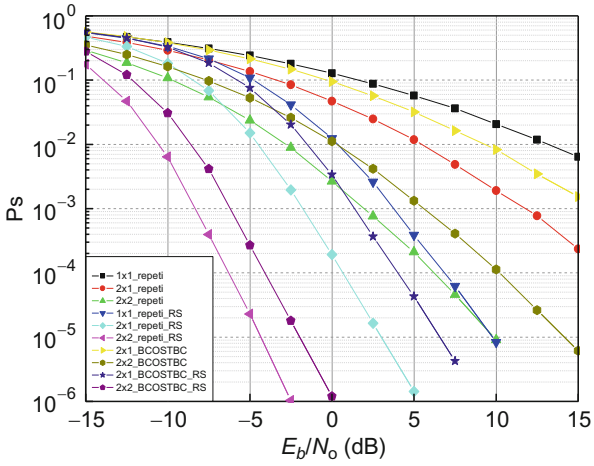


Fig. 20.63 Word-error rate of BCOSTBC code and repetitive code after concatenated RS (15, 7) when SI = 1

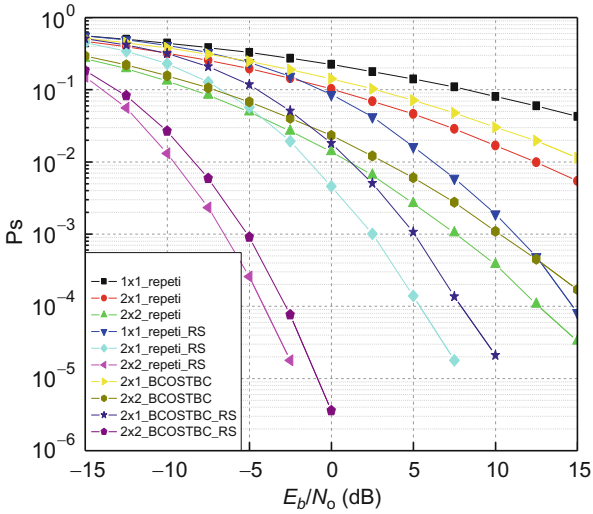


Fig. 20.64 Error rate of BCOSTBC code and repetitive code after concatenated RS (15, 7) when SI = 0.4

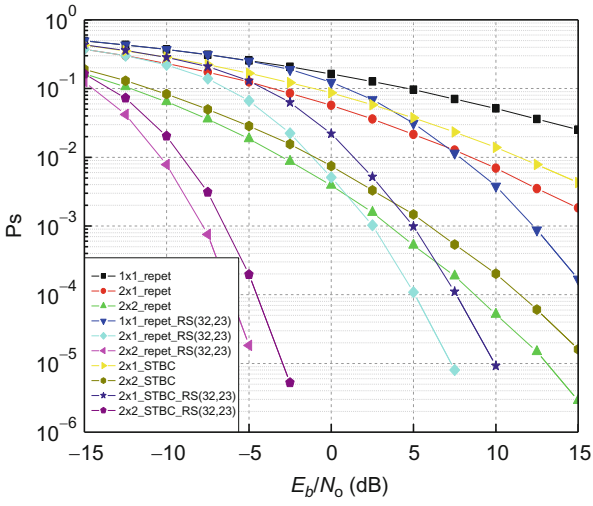


Fig. 20.65 When SI = 1, the error rate of BCOSTBC and repetitive codes after concatenated RS (32, 23) code

1. After concatenating the RS codes, the error rates of the BCOSTBC and repeat codes decrease significantly. When $P_s = 2 \times 10^{-3}$, in the 2×1 system, the BCOSTBC code after a concatenated RS (15, 7) achieves a 13.5-dB coding gain. After concatenating an RS (15, 7), a code gain of 12.5 dB is obtained. When $P_s = 2 \times 10^{-4}$, in the 2×2 system, the BCOSTBC code after a cascaded RS (15, 7) achieves a 13-dB coding gain. After concatenating RS (15, 7), a code gain of 12 dB is obtained. This shows that concatenated codes can significantly improve system performance.

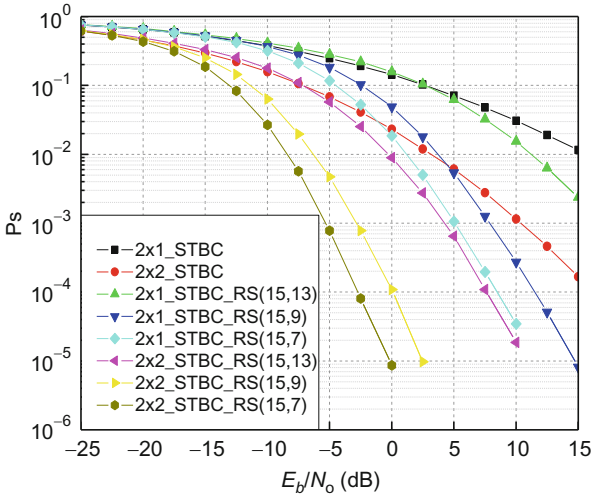


Fig. 20.66 Word-error rate of BCOSTBC code after concatenating different RS codes when SI = 1

2. In the same system, the performance gap between the two space–time codes before and after concatenation is significantly reduced. When $P_s = 2 \times 10^{-3}$, in the 2 × 1 system, the signal-to-noise ratio differences between the BCOSTBC code and the repeat code before and after concatenation are 4 dB and 3 dB, respectively. When $P_s = 2 \times 10^{-4}$, in the 2 × 2 system, the signal-to-noise ratio difference between the first two codes of the cascade is 4 dB, and that between the two cascade codes is 2 dB. This shows that in the BCOSTBC code, the error codes mainly originate from paired errors rather than random errors.
3. When the number of transmit apertures is the same, with an increase in the number of receive apertures, the bit-error performance gap between the two codes decreases significantly after concatenation. When $P_s = 2 \times 10^{-4}$, in the 2 × 1 and 2 × 2 systems, the signal-to-noise ratio differences between the two codes after concatenation are 3.3 dB and 2.2 dB, respectively. This indicates that the effect of the receiving diversity is better than that of the sending diversity.

Figure 20.64 shows the bit-error rate of the BCOSTBC and repeat codes after the concatenated RS (15, 7) code when SI = 1, also using 16-PPM modulation. It can be seen from the figure that the rule reflected in Fig. 20.64 is the same as that in Fig. 20.63, except that the improvement of the signal-to-noise ratio is greater at this time.

When $P_s = 2 \times 10^{-4}$, in the 2 × 1 and 2 × 2 systems, the signal-to-noise ratio differences between the two codes after concatenation are 2.5 dB and 1.2 dB, respectively. It can be seen that the performances of the two cascaded space–time codes are more similar, indicating that the cascaded space–time codes are more suitable for strong turbulence channels. In addition, the performance of the concatenated BCOSTBC code was slightly worse than that of the concatenated repetitive code, and the latter was more convenient for interleaving and concatenation.

Figure 20.65 shows the word-error rate of the BCOSTBC and repeat codes after concatenating the RS (32, 23) code when SI = 1. To facilitate the matching between symbols, 32-PPM modulation is adopted. Compared with Fig. 20.64, the coding gain of the last two space–time codes of the cascaded RS (15, 7) is greater than that of the last two space–time codes of the cascaded RS (32, 23) under a low SNR.

With the increase of SNR, the coding gain of the last two space–time codes of concatenated RS (32, 23) code increases dramatically. When the SNR is large, the coding gain of the last two space–time codes of the concatenated RS (15, 7) code is small. This shows that the error performance of the short code is better than that of the long code at a low SNR. However, with an increase in SNR, the high error-correction ability when using long codes can be brought into play.

Figure 20.66 shows the word-error rate of the BCOSTBC and repeat codes after concatenating the equal-length RS code when SI = 1. To facilitate matching between symbols, 16-PPM modulation was adopted. The following can be seen from the curve in the figure:

1. When the word-error rate $P_s \leq 10^{-3}$, the E_b/N_o requirements of various concatenated-code schemes are far less than those of non-concatenated ones. It can be seen that the concatenation of the RS code and space–time block code has obvious advantages.
2. The curves in the figure intersect. Before this intersection, the word-error rate decreases slowly with an increase in the SNR. Compared to concatenated codes with higher efficiency, concatenated space–time codes with lower efficiency exhibit poor bit-error performance. This is because the benefits of error correction cannot compensate for the impact of a higher bit-error rate caused by coding. After this intersection, the word-error rate curve rapidly becomes steeper. Compared to concatenated codes with higher efficiency, concatenated space–time codes with lower efficiency can obtain a higher coding gain. Therefore, an intersection is a turning point for codes with low efficiency.

Cascaded Layered Space–Time Codes

Cascaded Layered Space–Time Code Model

Figure 20.67 shows the basic structure of the LDPC-BLAST system. After information c is sent by the source, we encode it with LDPC to change it into a codeword x with a better code distance. After the serial-to-parallel conversion of the coded codeword x, layered space–time coding is performed, and the codeword is modulated using PPM modulation and then transmitted through an array of M lasers.

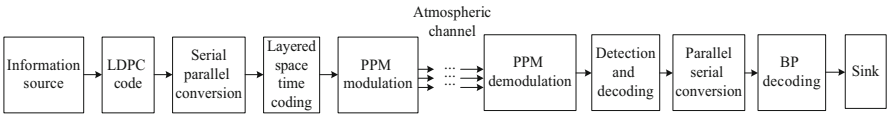


Fig. 20.67 LDPC-BLAST system model

At the receiving end, N detectors are used to receive signals, and PPM demodulation is performed on the received signals. The demodulated signals are detected and decoded to obtain the codeword. After a parallel-to-serial conversion, codeword y is obtained, and then the belief propagation (BP) decoding algorithm is applied to the codeword to obtain the final information c. Finally, we compare information c from the source with information c obtained at the receiver, calculate the corresponding bit-error rate, and judge the performance of the system according to the bit-error rate.

Analysis of Cascaded Layered Space–Time Codes

The maximum-likelihood, zero-forcing, and minimum mean-squared error detection algorithms were used to simulate the LDPC+BLAST cascade system with the D-BLAST and T-BLAST codes. The simulation conditions are as follows:

1. We assume the total power E_s is unchanged, and the value is 1 (normalized).
2. We assume that the channel characteristics and noise can be accurately estimated by the receiver.
3. The photoelectric-conversion efficiency is 0.6.
4. Under weak turbulence, the atmospheric channel follows a lognormal distribution, and the scintillation factor SI is 0.6. Under medium–strong and strong turbulence, the atmospheric channel follows a gamma–gamma distribution, and the scintillation factors SI are 1 and 3, respectively.
5. The system adopts pulse-position modulation, namely, 4-PPM.
6. LDPC code with binary domain rules is selected for coding, with a code rate of 0.5 and a code length of 2016 bits. A probabilistic BP decoding algorithm is adopted and the number of iterations is 100.
7. Antenna pairs 2×2, 2×4, 4×4, and 4×6 are used as examples.

Simulation of the LDPC+D-BLAST System

Figures 20.68 and 20.69 show the system simulation diagram when the scintillation SI $= 0.6$, that is, the curve of the bit-error rate and the signal-to-noise ratio of the

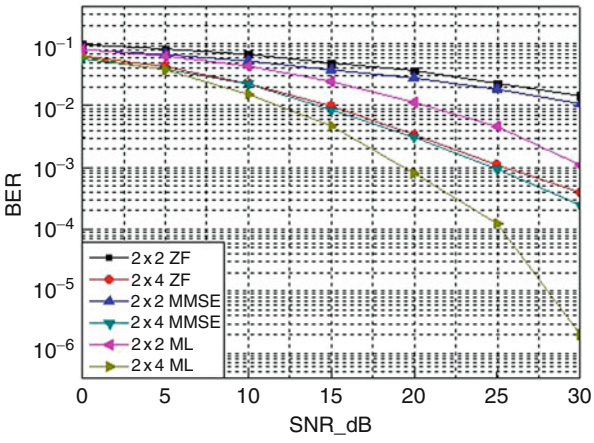


Fig. 20.68 Simulation diagram of 2×2 and 2×4 LDPC+D-BLAST systems (SI $= 0.6$)

LDPC+D-BLAST (LDPC code and D-BLAST cascade) system in weak turbulence. It can be seen that the bit-error rate of the system with four transmitters and multiple receivers is 10^{-6} when the SNR is 25 dB, while that of the system with two transmitters and multiple receivers can reach 10^{-6} only when the SNR is increased to 30 dB. Compared to the D-BLAST system without LDPC coding, the LDPC+D-BLAST system has better performance and lower signal-to-noise ratio requirements in the case of weak turbulence.

Figures 20.70 and 20.71 show the system simulation results when the scintillation index SI = 1, that is, the curve of the bit-error rate and the signal-to-noise ratio of the LDPC+D-BLAST system in the case of medium and strong turbulence. It can be observed that the enhanced SI has a greater impact on the bit-error rate. The LDPC+D-BLAST system needs to increase the system's signal-to-noise ratio

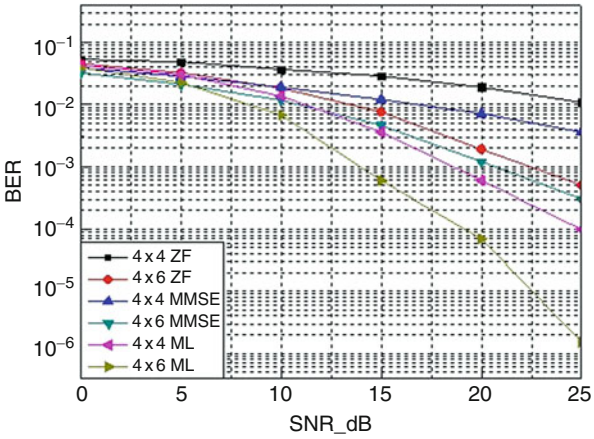


Fig. 20.69 Simulation diagram of 4 × 4 and 4 × 6 LDPC+D-BLAST systems (SI = 0.6)

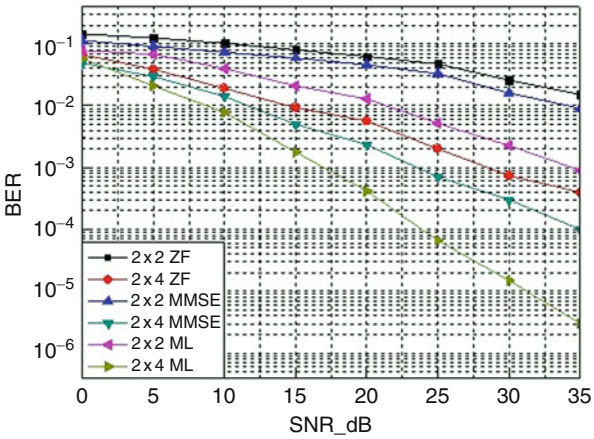


Fig. 20.70 Simulation diagram of 2 × 2 and 2 × 4 LDPC+D-BLAST systems (SI = 1)

to obtain the bit-error rate in a weak turbulence situation; that is, when the signal-to-noise ratio of the two-transmitter multi-receiver system is 35 dB, the bit-error rate can be close to 10^{-6}. This increases the system's transmitter-power requirements.

Figures 20.72 and 20.73 show the system simulation results when SI = 3, that is, the curve of the bit-error rate and signal-to-noise ratio of the LDPC+D-BLAST system in strong turbulence. It can be seen that when the signal-to-noise ratio is the same, the bit-error rate of the system in strong turbulence is much higher than that in weak turbulence (when the SNR is 25 dB, the bit-error rate in weak turbulence is 10^{-6}, and the bit-error rate in strong turbulence is 10^{-4}). In the case of strong turbulence, it is necessary to increase the transmission power at the transmitter and increase the signal-to-noise ratio to 40 dB to reduce the bit-error rate to nearly 10^{-6}.

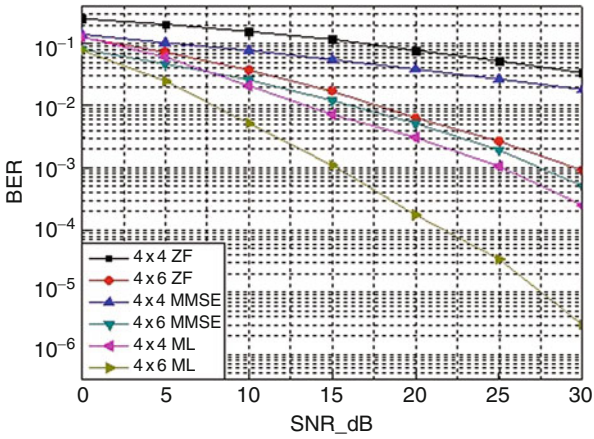


Fig. 20.71 Simulation diagram of 4 × 4 and 4 × 6 LDPC+D-BLAST systems (SI = 1)

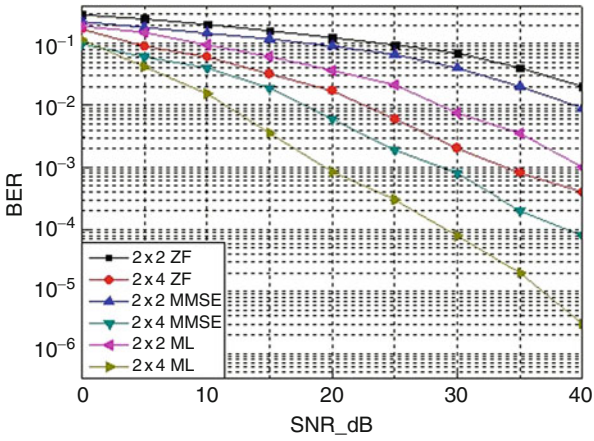


Fig. 20.72 Simulation diagram of 2 × 2 and 2 × 4 LDPC+D-BLAST systems (SI = 3)

LDPC+T-BLAST System

Figures 20.74 and 20.75 show the system simulation results when SI = 0.6, that is, the curve of the bit-error rate and signal-to-noise ratio of the LDPC+T-BLAST system in weak turbulence. Similar to the LDPC+D-BLAST system, under weak turbulence, the bit-error rate of the system with two transmitters and multiple receivers is close to 10^{-6} when the signal-to-noise ratio is 30 dB. This is slightly lower than that of the LDPC+D-BLAST system because the performance of diagonal layered space–time coding is better than that of spiral layered space–time coding. However, in a four-transmitter multi-receiver system, an SNR of only 25 dB can achieve a bit-error rate close to 10^{-6}. This indicates that the system performance can be significantly improved by increasing the number of transceiver antennas.

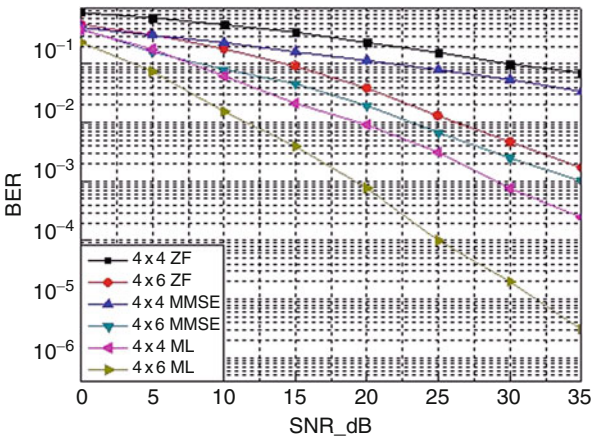


Fig. 20.73 Simulation diagram of 4 × 4 and 4 × 6 LDPC+D-BLAST systems (SI = 3)

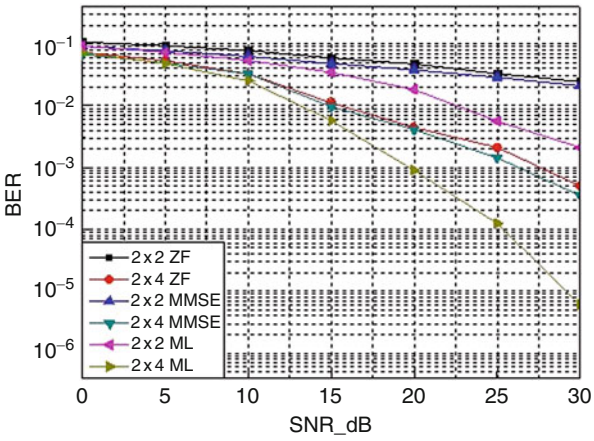


Fig. 20.74 Simulation diagram of 2 × 2 and 2 × 4 LDPC+T-BLAST systems (SI = 0.6)

Figures 20.76 and 20.77 show the system simulation results when SI = 1, that is, the curve of the bit-error rate and the signal-to-noise ratio of the LDPC+T-BLAST system in the case of medium and strong turbulence. It can be seen that the atmospheric-turbulence intensity affects the system performance to a certain extent. In a two-transmitter multi-receiver system, it is necessary to increase the signal-to-noise ratio to obtain the minimum bit-error rate under weak turbulence. The bit-error rate of the system is close to 10^{-6} when the SNR is 35 dB in moderate and strong turbulence. The bit-error rate also increases in the four-transmitter multi-receiver system. When the SNR increases to 30 dB, the bit-error rate decreases to approximately 10^{-6}.

Figures 20.78 and 20.79 show the system simulation results when SI = 3, that is, the curve of the bit-error rate and signal-to-noise ratio of the LDPC+T-BLAST system in strong turbulence. It can be observed that the performance of the system

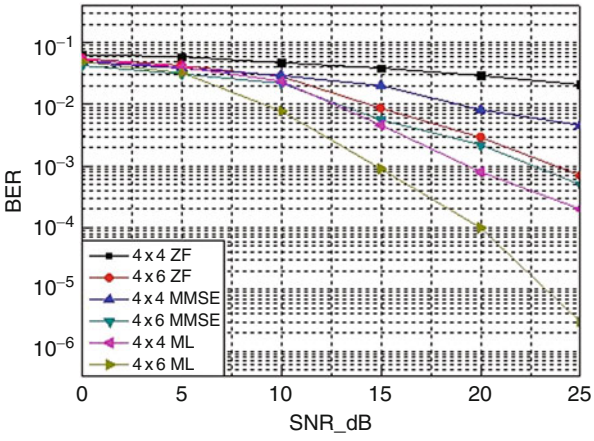


Fig. 20.75 Simulation diagram of 4 × 4 and 4 × 6 LDPC+T-BLAST systems (SI = 0.6)

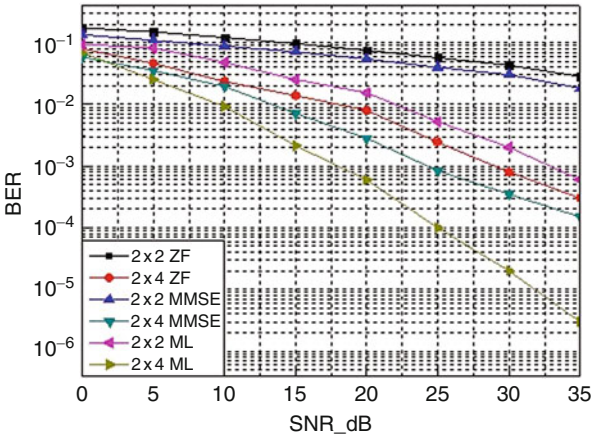


Fig. 20.76 Simulation diagram of 2 × 2 and 2 × 4 LDPC+T-BLAST systems (SI = 1)

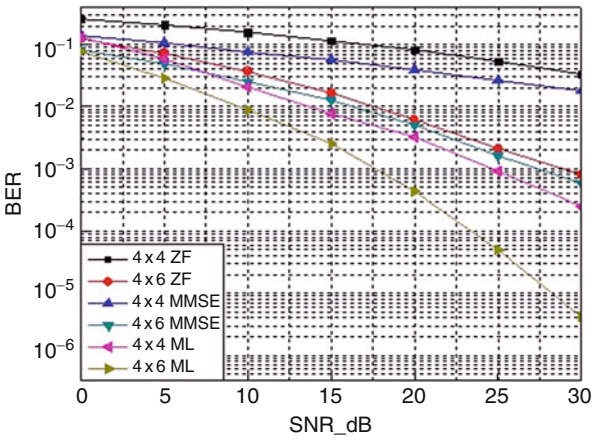


Fig. 20.77 Simulation diagram of 4×4 and 4×6 LDPC+T-BLAST systems (SI = 1)

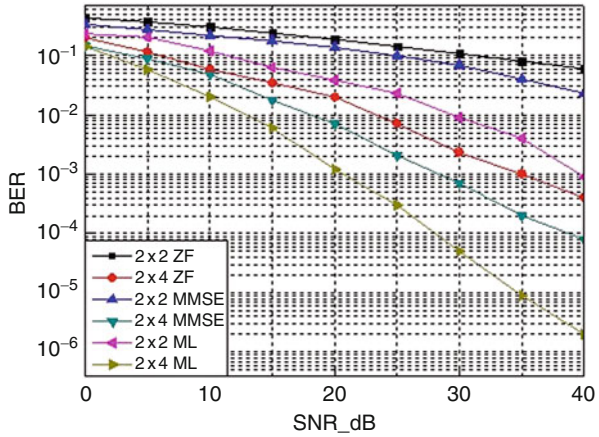


Fig. 20.78 Simulation diagram of 2×2 and 2×4 LDPC+T-BLAST systems (SI = 3)

worsens in the case of strong turbulence, and the bit-error rate value increases under the same SNR. When the signal-to-noise ratio is 25 dB, the bit-error rate of the two-transmitter multi-receiver system increases from 10^{-4} to 10^{-3}, while the bit-error rate of the 4×4 system increases from 10^{-5} to 10^{-4}.

To reduce the effect of strong turbulence on the system, it is necessary to increase the system transmission power, that is, to increase the signal-to-noise ratio. When the signal-to-noise ratio is increased to 40 dB in a two-transmitter multi-receiver system, the system's bit-error rate is close to 10^{-6}, while the signal-to-noise ratio of a four-transmitter multi-receiver system only needs to be increased to 35 dB.

Although the LDPC code increases the complexity of the system, the error rate of the LDPC+BLAST system decreases significantly. This shows that the LDPC code can improve the bit-error rate of the system, making it better than a single BLAST

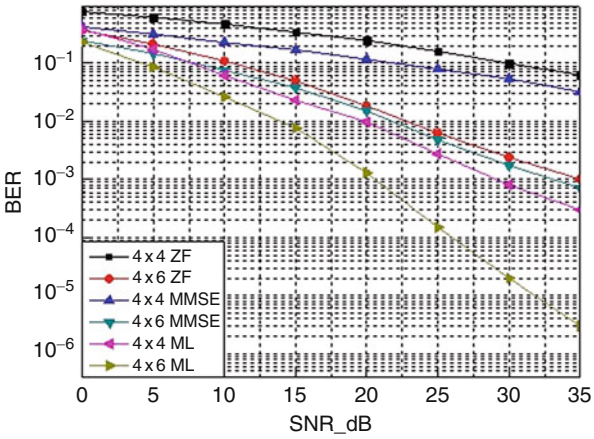


Fig. 20.79 Simulation diagram of 4×4 and 4×6 LDPC+T-BLAST systems (SI = 3)

system. When the signal-to-noise ratio is 30 dB in weak turbulence, the bit-error rate of the two-transmitter multi-receiver system decreases from 10^{-4} to 10^{-6}, indicating that the system with LDPC coding can better overcome the interference of noise in the system and improve the reliability of the system transmission information.

Cascade coding is an efficient coding method that has been widely used in recent years to enable the system error correction to approach the Shannon limit [127]. We constructed an RS-code concatenated space–time code and an LDPC-code concatenated space–time code. In contrast to [127, 128], we expanded the system to 4×4 and 4×6. Cascaded space–time codes can improve the bit-error rate of the system in different atmospheric-turbulence environments. Reducing the operational complexity of concatenated coding and improving operational efficiency are subjects for future study [129–132].

References

1. Ma Liang. Smart Antenna in Wireless Communication - IS95 and 3G CDMA Applications [M]. Beijing: Machinery Industry Press, 2002:20–31.
2. Huang Hao-xue, Wu Si-liang. An Algorithm for Estimating DOA and Doppler Frequency of Signals Incident on Uniform Circular Array [J]. Acta Electronica Sinica, 2001, 29(5): 619–624.
3. Khalighi, M.A, Brossier, J.-M, Jourdain, G.V, Raoof, K. Water filling capacity of Rayleigh MIMO channels [J]. Personal, Indoor and Mobile Radio Communications, 12th IEEE International Symposium on. Oct, 2001, 1:156–158.
4. Gao Jie. Applying MIMO Technology in the Ultra Wide Band System [D]. Xi'an: Xidian University, 2008:10–11
5. Stéphanie S, Anne J-V, Jean-Pierre C. Soft decision LDPC decoding over chi-square based optical channels [J]. Journal of Lightwave Technology, 2009, 27(16):3540–3545.
6. Laurent B, Planche G. SILEX overview after flight terminals campaigh [J]. SPIE, 1997, 2990(97): 10–26.
7. Fenner W R. Future trends in crosslink communications [J]. SPIE, 1993, 1866(01): 1–8.

8. Gotthard O. SILEX program status-a major milestone is reached [J]. SPIE, 1997, 2990(07): 2–9.
9. Bondurant R S, Boroson D M. Overview of lasercom program at Lincoln laboratory [C]// Free-Space Laser Communication Technologies VII, Lexington, MA, USA, 20 April, 1995: 2–3.
10. D. Divsalar R. M. Gagliardi, J. H. Yuen. PPM performance for Rees-solomon decoding over an optical-Rf relay link [J]. IEEE Transactions on Communications, 1984, 32(3):302–305.
11. Koh Y. T. DF. Interleaved concatenated coding for the turbulent atmospheric direct detection optical communication channel [J]. IEEE Transactions on Communications, 1989, 37(6): 648–651.
12. H. Park, J R B. Trellis-coded multiple-pulse-position modulation for wireless infrared communications [J]. IEEE Transactions on Communications, 2004, 52(4):643–651.
13. K Seki KM, A. Katayama, S. Suzuki, N. Shinohara, M. Nakabayashi. Single-chip FEC code using a concatenated BCH code for 10 Gb/s long-haul optical communication systems [C]. IEEE 2003 Custom in Integrated Circuits Conference, 2003: 279–280.
14. Yi Cai JMM, Tülay Adalı, etc. On Turbo code decoder performance in optical-fiber communication systems with dominating ASE noise [J]. Journal of light wave technology, 2003, 21(3):727–734.
15. K. Chitra V C R. Turbo-coding of coherence multiplexed optical PPM CDMA system with balanced detection [J]. American Journal of Applied Sciences, 2007, 4(5):264–268.
16. Stephanie Sahuguede DF, Anne Julien-Vergonjanne, Georges Rodriguez, Jean-Pierre Cances PG. LDPC code design and performance analysis on OOK Chi-Square-Based optical channels [J]. IEEE photonics technology letters, 2009, 21(17):1190–1192.
17. Davidson F M K, Y. T. Interleaved convolution coding for the turbulent atmospheric optical communication channel [J]. IEEE Transactions on Communications, 1988, 36:993–1003.
18. Maurizio Magarini R-JE, Bert E. Basch, etc. Concatenated Coded Modulation for Optical Communications Systems [J]. IEEE photonics technology letters, 2010, 22(16): 1244–1246.
19. Djordjevic IB, Lei Xu. Ting Wang. Reverse concatenated coded modulation for high-speed optical communication [J]. IEEE Photonics Journal, 2010, 2(6): 1034–1039.
20. Wilson K E. Results of the compensated Earth-Moon-Earth retro-reflector laser link (CEMERLL) experiment [R]. TDA Progress Report 42–131, 1997:1–13.
21. Kim I I. Scintillation reduction using multiple transmitters [J]. SPIE, 1997, 2990(09): 102–113.
22. Kim I I, Woodbridge E L, Chan V J, et al. Scintillation measurements performed during the limited-visibility lasercom experiment [J]. Proceedings of SPIE-The International Society for Optical Engineering, 1998, 3266(05): 209–223.
23. Kim I I, Mitchell M, Korevaar E J. Measurement of scintillation for free-space laser communication at 785 nm and 1550 nm [J]. Proc Spie, 1999, 3850(05): 49–62.
24. Kim I I, Koontz J, Hakakha H, et al. Measurement of scintillation and link margin for the TerraLink laser communication system [J]. Proc Spie, 1998, 3232(05): 100–118.
25. Wilson K E. Results from Phase-l and Phase-2 GOLD experiments [R]. TDA Progress Report 42–133. 1997.
26. Strohbehn J W. Laser beam propagation in the atmosphere [M]. New York: Spring-Verlag Berlin Heidelberg, 1978.
27. Tatarskii V I. Wave propagation in a turbulent medium [M]. New York: McGraw, 1961.
28. Ibrahim M M, Ibrahim A M. Performance analysis of optical receivers with space diversity reception [J]. IEEE Proc Commun, 1996, 143(6):369–372.
29. Nykolak G, Szajowski P F, Cashion A, et al. 40-Gb/s DWDM free-space optical transmission link over 4.4 km [J]. Proceedings of SPIE-The International Society for Optical Engineering, 2000, 3932(01): 16–23.
30. Biswas A, Wright M W. Mountain-Top-to-Mountain-Top Optical Link Demonstration [R]. Part I. IPN Progress Report, 2002: 42–149.

31. Vilnrotter V, Mecherle G S, Lau C W, et al. Real-time combining of optical array signals [C]// Lasers & Applications in Science & Engineering, San Jose, California, 28 March, 2006: 610508.1–610508.11.
32. Vilnrotter V A, Lau C W, Andrews K, et al. Two-Element Optical Array Receiver Concept Demonstration [J]. Proceedings of SPIE-The International Society for Optical Engineering, 2005, 5712(06): 225–239.
33. V Vilnrotter, C. W. Lau, M. Srinivasan, etc. Optical array receiver for communication through atmospheric turbulence [J]. IEEE Journal of Lightwave Technology, 2005, 23(4):1664–1675.
34. Vilnrotter V, Srinivasan M. Optical communications through atmospheric turbulence using photodetector arrays [J]. Proc Spie, 2001, 4272(06): 282–292.
35. S. M. Haas, etc. Capacity of Optical Wireless Communication[J]. IEEE Journal on Selected Areas in Communications, 2001, 21(8):1346–1357.
36. Ma Dongtang, Wei Jibo, Zhuang Zhaowen. Space laser communication and its application [J]. Semiconductor Optoelectronics, 2003, 24(2):139–144.
37. Haas S M, Shapiro J H, Tarokh V. Space-time codes for wireless optical channels [C]// IEEE International Symposium on Information Theory, Washington, DC, USA, 29 June, 2001: 244.
38. X. Zhu, etc. Mitigation of turbulence-induced scintillation noise in free-space optical links using temporal-domain detection techniques [J]. IEEE Photonics Technology Letters, 2003, 15(4):623–629.
39. X. Zhu, etc. Free-space optical communication through atmospheric turbulence channels [J]. IEEE Transactions on Communications, 2002, 50(8):1293–1300.
40. M. K. Simon, etc. Alamouti-type space-time coding for free-space optical communication with direct detection [R]. IPN Progress Report 42–156, 2003; 2003: 1–9.
41. Alqudah Y, Somani A K, Zhang Z, et al. Orthogonal spatial coding in indoor wireless optical link reducing power and bandwidth requirements [J]. International Society for Optics and Photonics, 2003, 5285(07): 237–245.
42. Feng P, Yu S Y. Scintillation characterization of multiple transmitters for ground-to-satellite laser communication [J]. 2004, 5640(06): 448–454
43. Wilson SG, Brandt-Pearce M, Cao Q. Optical MIMO transmission with Q-ary PPM for atmosphere channel [J]. IEEE Transactions Communications, 2004, 53(8):1090–1094.
44. Maïté Brandt-Pearce, Stephen Wilson, Qianling Cao, etc. Code design for optical MIMO systems over fading channels [C]. Proceedings of the 38th Asilomar conference on signals, systems & computers, Monterey, CA, 2004: 871–875.
45. M. K. Simon, Vilnrotter V. Alamouti-type Space-time coding for free-space optical communication with direct detection [J]. IEEE Transactions on wireless communications, 2005, 4(1): 35–39.
46. Anguita J A, Neifeld M A, Vasic B V. Multi-beam space-time coded systems for optical atmospheric channels [J]. International Society for Optics and Photonics, 2006, 6304(02): 63041B.1–63041B.9.
47. Jaime A. Anguita, Mark A. Neifeld, Vasic B V. Spatial correlation and irradiance statistics in a multiple-beam terrestrial free-space optical communication link [J]. Applied Optics, 2007, 46(26): 6561–6571.
48. S. Mohammad Navidpour, Murat Uysal, Kavehrad M. BER performance of free-space optical transmission with spatial diversity [J]. IEEE Transactions on Wireless Communications, 2007, 6(8): 2813–2821.
49. Qianling Cao, Maïté Brandt-Pearce, Wilson S G. Free space optical MIMO system using PPM modulation and a single optical Amplifier [J]. Communications and networking in China, 2007: 508–512.
50. García-Zambrana A. Error rate performance for STBC in free space optical communications through strong atmospheric turbulence [J]. IEEE communication letters, 2007, 11(5):390–392.
51. Cvije tic Neda, Wilson Stephen G., Brandt-Pearce Maite. Performance bounds for free-space optical MIMO systems with APD receivers in atmospheric turbulence [J]. IEEE Journal on Selected Areas in Communications, 2008, 26(3):3–11.

52. Chakraborty K. Outage capacity of MIMO Poisson Fading Channel [J]. IEEE transactions on information theory, 2008, 54(11):4887–4907.
53. Nick Letzepis, Ian Holland, Cowley W. The Gaussian free space optical MIMO channel with Q-ary pluse position modulation [J]. IEEE transactions on wireless communications, 2008, 7(5): 1744–1753.
54. Majid Safari, Uysal M. Do we really need OSTBCs for free-space optical communication with direct detection [J]. IEEE transactions on wireless communications, 2008, 7(11):4445–4448.
55. Zaidi S, Hafeez M. Cross layer design for orthogonal space time block coded optical MIMO systems [C]// Ifip International Conference on Wireless & Optical Communications Networks, Surabaya, Indonesia, 07 May, 2008: 1–5.
56. Georgia Ntogari, T.K., Thomas Sphicopoulos. Performance analysis of space time block coding techniques for indoor optical wireless systems [J]. IEEE Journal on Selected Areas in Communications, 2009. 27(9): 1545–1552.
57. Ehsan Bayaki, R.S. On space–time coding for free–Space optical systems [J]. IEEE transactions on wireless communications, 2010. 58(1): 58–62
58. Antonio García-Zambrana, Carmen Castillo-Vázquez, Beatriz Castillo-Vázquez, Space-time trellis coding with transmit laser selection for FSO links over strong atmospheric turbulence channels [J]. OPTICS EXPRESS, 2010. 18(6): 5356–5366.
59. Qiang Xiwen, Wang Tieliang, Wu Naiqing. Multi-Laser-Beams Atmospheric Propagation [J]. OPTOELECFRONIC TECHNOLOGY, 1999, 19(3):167–172.
60. Liu Weihui, Wu Jian. Influence of Space between Atmospheric Channels and Beams' Number on Scintillation [J]. Optics & Optoelectronic Technology, 2003, 1(4): 15–19.
61. Liu Weihui, Wu Jian. Influence of multiple partially coherent beams through strong turbulence on light intensity scintillation [J]. Opto-electronic Engineering, 2004, 31(4):5–8.
62. Ma Dong-tang. Study on Multiple-beam Transmission and Reception Techniques for Atmospheric Laser Communication [D]. Changsha: National University of Defense Science and Technology, 2004.
63. Ma Dongtang, Hua Weiling, Wei Jibo. Analysis of the Laser Diodes Temperature Field in Multiple-beam Atmospheric Laser Communications [J]. Journal of Optoelectronics.laser, 2004, 15(12): 1449–1451.
64. Ma Dongtang, Wei Jibo, Zhuang Zhaowen. Performance Evaluation and Channel Modeling of Multiple-Beam Propagation for Atmospheric Laser Communication [J]. Acta Optica Sinica, 2004, 24(8): 1020–1028.
65. Chen Gang, Dong Zuo ren, Geng Jian xin, et al. 155/622 Mb/s Multiple Transmitter Laser Communication Systems [J]. Chinese Journal of Lasers, 2004, 31(5): 583–587.
66. Deng Daizhu, Guo Huafu. Application of Multiple Transmitters in Atmospheric Laser Communication [J]. Optics & Optoelectronic Technology, 2005, 3(4):13–15.
67. Wu Changqi, Pan Jimin, Yu Rongjin. Atmospheric optical communication with optimal selection diversity [J]. Infrared and Laser Engineering, 2005, 34(1): 114–130.
68. He Jian, Hu Yanjun. Performance analysis of wireless optical MIMO communications based on SVD [J]. Journal of Anhui University, 2007, 31(3): 34–37.
69. Zheng Zheng, Li Ning. Diversity detection technology in laser communication [J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(3):280–284.
70. Tian Zairong, huaibei Z. Structure and performance of space-time block code. Wireless Communications [J]. Networking and Mobile Computing, 2006: 1–4.
71. Chen Jianwen. Research on Laser's Transmission Channels and Properties Improvement of Optical Communication in the Atmosphere [D]. Wuhan: Huazhong University of Science and Technology, 2007.
72. Deng Tianping. Research on the Key Technology of FSO System [D]. Wuhan: Huazhong University of Science and Technology, 2007.
73. Zhang Hongliang. The Research of the Space-Time Code Basede on Launch Diversity for Atomospheric Laser Communication [D]. Changchun: Changchun University of Technology, 2008.

74. Chen Chunyi. Study on Mechanism and Mitigation Technology of Atmospheric Effects in Optical Wireless Communications [D]. Changchun: Changchun University of Technology, 2009.
75. Huang Haoxue, WU Siliang. An Algorithm for Estimating DOA and Doppler Frequency of Signals Incident on Uniform Circular Array [J]. Acta Electronica Sinica. 2001, 29(5):619–620.
76. Khalighi M A, Brossier J M, Jourdain G V, et al. Water filling capacity of Rayleigh MIMO channels [C]// IEEE International Symposium on Personal, San Diego, CA, USA, 06 August, 2002: 156–158.
77. Gao Jie. Applying MIMO Technology in the Ultra Wide Band System [D]. Xi'an: Xidian Univer, 2008:10–11
78. Li Ting, The Research on Layered Space-time Code for FSO System [D], Xi'an: Xi'an University of Technology, 2012
79. Wang Huiqin, Ke Xizheng. Free Space Optical Communication Based on Vertical Bell Labs Layered Space-Time [J]. Chinese Journal of Lasers 2008, 35(6):874–878.
80. Ke Huanhuan, The Research on Vertical Bell Labs Layered Space-time for FSO MIMO System [D], Xi'an: Xian University of Technology, 2011
81. Ge Ziye, Research and Implementation of FSO MIMO Vertical Layered Space-Time Code [D], Xi'an: Xi'an University of Technology, 2010
82. Ke Xizheng, Wang Huiqin, Free-space MIMO optical communication system based on vertical layered space-time coding [P], Chinese patent: ZL200810018222, 2012.08
83. Wang Huiqin, Research of space-time code in atmosphere laser communication [D], Xi'an: Xi'an University of Technology, 2011
84. He Hua, Study on the Key Technology of Wireless Ultraviolet Communication and Ultraviolet Networking [D], Xi'an: Xi'an University of Technology, 2011
85. Gao Haitao, The Research on Space-time Trellis code for Woc MIMO ststem [D], Xi'an: Xi'an University of Technology, 2011
86. Ge Ziye, Research and Implementation of FSO MIMO Vertical Layered Space-Time Code [D], Xi'an: Xi'an University of Technology, 2010
87. Zhao Li, Research on the Key Techniques of FSO-OFDM [D], Xi'an: Xi'an University of Technology, 2009
88. Luo Tao, Le Guangxin. Principle and application of multi-antenna wireless communication [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2005.
89. M. K. Simon, Vilnrotter V. Alamouti-type Space-time coding for free-space optical communication with direct detection [J]. IEEE Transactions on wireless communications, 2005, 4(1): 35–39.
90. Vahid Tarokh, Nambi Seshadri, and Calderbank A R. Space-Time Codes for High Data Rate Wireless Communications: Performance Criterion and Code Construction [J]. IEEE Trans. Inform. Theory, 1998–03, 44:744–765
91. V. Tarokh, N. Seshadri and A.R. Calderbank, Space-time codes for high data rate wirless communication performance criterion and code construction [J], IEEE Trans. Inform. Theory, Mar. 1998,44(2):744–765.
92. Tarokh V, Jafarkhani H.A Differential Detection Scheme for Transmit Diversity [J]. IEEE Journal on Selected Areas in com, 2000. 18(O7):1169–1174.
93. Alamouti S M.A Simple Transmit Diversity Scheme for Wireless Communication [J]. IEEE J. Se1.Areas Commun. 1998, 16(O8):1451–1458
94. Tarokh V, Jafarkhani 1t. Multiple Transmit Antenna Differential Detection from Generalized Orthogonal Designs [J]. IEEE Trans. Inf. Theory, 2001. 47(06):2626–2631.
95. Hochwald M B. Differential Unitary Space-Time Modulation [J]. IEEE Transactins on Com. 2000, 48(12):2041–2052.
96. Hughes B L. Differential Space-Time Modulation [J]. IEEE Transactions on Information Theory. 2000, 46(O7):2567–2578.

97. M. K. Simon, Vilnrotter V. Alamouti-type Space-time coding for free-space optical communication with direct detection [J]. IEEE Transactions on wireless communications, 2005, 4(1): 35–39.
98. Brian L Hughes. Differential Space-Time Modulation [J]. IEEE Trans. on Information Theory, 2000–11, 46(7):2567–2578
99. Hochwald B M, Marzetta T L. Systematic Design of Unitary Space-Time Constellations [J]. Information Theory IEEE Transactions on, 2000, 46(6): 1962–1973.
100. Bertrand M. Hochwald and Thomas L. Marzetta, Unitary Space-Time Modulation for Multiple-Antenna Communications in Rayleigh Flat Fading [J]. IEEE Trans. Information Theory, 2000–03, 46(2):543–564
101. Wolniansky P W. V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel [C]// Ursi International Symposium on Signals, Pisa, Italy, 02 October, 1998: 295–300.
102. Li Ying, Li Jia, Wang Xinmei. An improved detection algorithm of layered space-time codes [J]. Journal of China Institute of Communications, 2003, 24(3):113–118.
103. Maïté Brandt-Pearce, Stephen Wilson, Qianling Cao, etc. Code design for optical MIMO systems over fading channels [C]. Proceedings of the 38th Asilomar conference on signals, systems & computers, Monterey, CA, 2004: 871–875.
104. M. K. Simon, Vilnrotter V. Alamouti-type Space-time coding for free-space optical communication with direct detection [J]. IEEE Transactions on wireless communications, 2005, 4(1): 35–39.
105. Syed Ali Raza Zaidi, afees M. Cross Layer Design for Orthogonal Space Time block Coded Optical MIMO Systems [C]. WOCN ’08. Wireless and optical communications networks, 2008: 1–5.
106. Wilson S G, Brandt-Pearce M, Cao Q, Baedke M. Optical Repetition MIMO Transmission With Multipulse PPM [J]. IEEE Journal on selected areas in communications, 2005, 23(9): 1901–1910.
107. Luo Tao, Le Guangxin. Principle and application of multi-antenna wireless communication [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2005.
108. Wang Huiqin Ke Xizheng. Hybrid Space-Time Code for Free Space Optical Communication [J]. Acta Optica Sinica, 2009, 29(1):132–137.
109. Li Ting, The Research on Layered Space-time Code for FSO System [D], Xi’an: Xi’an University of Technology, 2012
110. Qiu Shaofeng, Ren Wei, Li Jibi. Study of optical fiber multi-input multi-output system based on V-BLAST [J]. Optical Communication Technology,2011,01:11–13.
111. Li Jibi, Ren Wei, Qiu Shaofeng. Detection Algorithm in V-BLAST Optical MIMO System [J]. Semiconductor Optoelectronics, 2011, 32(01):96–104.
112. Zhao Guhao, Zhao Shanghong, Li Yongjun, et al. Free Space Optical Multiple-Input Multiple-Output Communication Based on Intensity Superposition Code [J]. Acta Optica Sinica, 2011, 31(07):62–67.
113. Xing Xuefeng, Li Hongzuo. Space Laser Communications Based on Quasi-Orthogonal Space-Time Block Coding [J]. Chinese Journal of Lasers, 2012, 39(05): 156–162.
114. Li Jinghua. Research of PPM-based Reed-solomon concatenate codes modulation in optical wireless communications [J]. Optical Communication Technology, 2012, 3(02): 30–32.
115. Pan Yan, Pang Weizheng, Wang Xuehuan, et al. The algorithm for differential concatenated decoding based on Turbo coding [J]. Applied Science and Technology, 2006, 33(06): 57–59.
116. Wang Hua, Xiao Jianming, Kuang Jingming. Performances of P TCM Concatenating with RS Codes [J]. Journal of Beijing Institute of Technology, 1999, 19(01): 67–71.
117. Wang Duanyi; Yue Guangxin. A New Concatenated Coding System Using Multilevel Inner Codes [J]. Journal of Beijing University of Posts and Telecommunication, 1994, 17(01): 1–8.
118. Li Ying, Guo Xudong, Wang Xinmei. Design of Generalized Recursive Space-Time Trellis Codes for Serial Concatenation [J]. Journal of Electronics & Information Technology, 2005, 27(11):54–58.

119. Forney. G.D., Coneatenated Codes, MIT Press, Cambridge, 1966.
120. Xu Jianwu, Wang Hongxing. A Cascade Space-time Coding Method Based on TPC and OSTBC in FSO Communication Systems [J]. Telecommunication Engineering,2012,52(10): 1596–1601.
121. Wang Huiqin, Cao Minghua. Concatenated Space-time Block Code in Atmosphere Laser Communication [J]. Opto-Electronic Engineering, 2010. 37(12):116–124.
122. C G J C, B C J. Error-Correction Coding for Digital Communications [M]. Plenum Press, 1981.
123. Yuan Dong feng, Zhang Lijun. A theoretical method for estimating the performance of concatenated codes in Rayleigh fading channel [J]. Journal of China Institute of Communications, 2000, 21(10): 86–89.
124. Wilson SG, Brandt-Pearce M, Cao Q. Optical MIMO transmission with Q-ary PPM for atmosphere channel [J]. IEEE Transactions Communications, 2004, 53(8):1090–1094.
125. Wilson SG, Brandt-Pearce M, Cao Q. Optical MIMO transmission with Q-ary PPM for atmosphere channel [J]. IEEE Transactions Communications, 2004, 53(8):1090–1094.
126. Xiao Haiwei1, Wang Yang, Zhang Jiliang, et al. Concatenated coding based on LDPC codes and super quasi-orthogonal STBC, Systems Engineering and Electronics, 2011,33(9): 2122–2125
127. Choi C S, Shin J H. Construction and performance analysis of thecombined system of STBC and LDPC code achieving data rate [J].Codingand Information Theory Lab, 2007,43(2): 1763–1768.
128. Futaki H, Ohtsuki T. LDPC-based space-time transmit diversity schemes with multiple transmit antennas [C]||Proc. ofthe57th IEEE Semiannual Vehicular Technology Conference, 2003:2589–2593.
129. Ke Xizheng, Chen Juan, Deng Lijun. Research progress of space-time code in Optical Wireless Communication (I) [J]. Infrared and Laser Engineering, 2013, 52(07):1882–1889
130. Ke Xizheng; Yuan Lei; Li Fang. Research progress of space-time code in Optical Wireless Communication (II) [J]. Infrared and Laser Engineering, 2013, 42(08):2137–2145.
131. Ke Xizheng; Chen Juan; Li Zheng. Research progress of space-time code in Optical Wireless Communication (III) [J]. Infrared and Laser Engineering, 2013, 42(09):2496–2504.
132. Ke Xizheng; Chen Juan; Chen Dan. Research progress of space-time code in Optical Wireless Communication (IV) [J]. Infrared and Laser Engineering, 2013, 42(10):2765–2771.

Experimental Study on Optical Wireless Coherent Communication

21

Contents

Introduction 694
Foreign Research Progress 694
Domestic Research Progress 697
Experiment on Coherent Detection in Radio Optical Communication 701
Coherent-Detection Techniques 701
Analysis of Coherent Detection Systems Under Atmospheric Turbulence 703
Analysis of a Coherent-Detection System Under the Influence of Channel Noise 709
Analysis of the Influence of Optical-Signal Polarization on Coherent Detection Systems 713
Performance Analysis of the Light Source on a Coherent Detection System 716
Performance Analysis of a Subcarrier-Modulated (SCM) Nonoptical Outlier-Detection System 718
Color Difference of a Dual-Wavelength Free-Space Optical Communication System 724
Coherent-Detection Performance Optimization 737
Influence of the Polarization State of Partially Coherent Light on the Sensitivity of Coherent Detection in Non-Kolmogorov Atmospheric Turbulence 737
Effect of the Optical Spatial Distribution on the Mixing Efficiency 750
Coherent Detection System 751
Mixing Efficiency of a Gaussian-Distributed Signal Light and Various Local Vibration Lights 754
Mixing Efficiency When the Signal Light Has an Airy Distribution with Varying Local Vibration Lights 757
Mixing Efficiency When the Signal Light Is a Vortex Beam with Varying Local Vibration Lights 762
Influence of Wavefront Distortion on Coherent Detection Performance 773
Theoretical Analysis 774
Numerical Analysis 779
Experimental Study 786
References 790

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_21

Abstract

Compared with direct detection, coherent detection has the advantages of a fast response speed, high detection sensitivity, and good filtering performance. It is the key technology for improving the receiving sensitivity of optical wireless communication systems (Xingzheng Ke, Ke Dong, Optical Wrieless Communication [M]. Springer Press, Beijing, 2022). This chapter summarizes the domestic and foreign research progress in the field of coherent detection in optical wireless communication systems, and introduces the research work of Xi'an University of Technology in coherent-detection technology. It mainly includes the effects of atmospheric-turbulence suppression, arrival-angle fluctuation, amplitude fluctuation, channel noise, optical-signal polarization, local vibration light power, and light-source linewidth on the performance of coherent-detection systems. It also describes the research progress on subcarrier-modulated nonoptical extraterritorial detection systems and a performance analysis of ideal and nonideal channels. Finally, the development directions of optical wireless coherent communication systems are prospected.

Keywords

Optical wireless communication · Coherent detection · Atmospheric turbulence · Wavefront aberration

Introduction

Compared to intensity modulation/direct-detection systems, coherent optical communication systems can have a gain in detection sensitivity of about 20 dB (up to 23 dB for zero-difference detection), which does not depend on the light used in the transmission, but rather on the coherent modulation used at the transmitter end and coherent techniques used for detection at the receiver end. It allows transmission over longer distances at the same transmit power and is more favorable to suppress atmospheric turbulence and channel fading [1]. This chapter describes coherent optical communication and its experiments.

Foreign Research Progress

Optical wireless coherence communication, as a supplement or extension of optical-fiber communication, breaks through the technical bottleneck of traditional communication. At present, the optical-fiber communication field is mature. To further promote the practical application of coherent communication and improve the sensitivity and transmission capacity of coherent detection, researchers have explored environmental factors, polarization technology, and other coherent-detection methods.

In the 1970s, Fink [3] gave a general equation for the signal-to-noise ratio (SNR) of a coherent receiver and showed that coherent detection was an effective method to obtain the quantum-detection limit in the spectral range.

In 1975, Cohen [4] analyzed the influence of wavefront deformability and optical-field distribution on heterodyne detection. He found that a heterodyne-detection system with a uniformly distributed signal-light amplitude was more likely to produce a serious wavefront mismatch than that with an Airy distributed signal light, and proposed a general equation for the signal-to-noise ratio of a coherent receiver.

In 1978, Takenaka et al. [5] studied the heterodyne efficiency of coherent detection. The communication quality could only be guaranteed when the value of the differential efficiency was greater than 0.8 in the system design. In addition, the signal-to-noise ratio of Gaussian-field optical heterodyne detection was studied.

In 1984, Tanaka et al. [6] studied the influence of background noise on the signal-to-noise ratio of a differential detection system with limited local light power. When the ratio of the signal power to local light power and the ratio of the background-noise power to local light power were both set to small values, the signal-to-noise ratio of the system was closer to the quantum noise limit.

In 1987, Tanaka and Ohta [7] analyzed the influence of the wavefront-inclination angle and the optical-axis offset on the differential efficiency when the signal light had a Gaussian distribution. When the signal light-front inclination angle was less than 10^{-4} and the optical-axis offset was less than one-fifth of the incident aperture, the heterodyne efficiency reached more than 60%.

In 1992, Iftekharuddin et al. [8] analyzed the advantages of using nondiffused beams in heterodyne detection. In this case, the maximum heterodyne efficiency was about 50% when the incident inclination angle of a Bezer beam was 1.4°.

In 1994, Kalman et al. [9] studied the performance of optical wireless coherent communication links using amplitude, phase, and frequency modulation. Compared with direct-detection links, optical wireless coherent communication links provided higher anti-interference ability.

In 1997, Chambers [10] analyzed the influence of the heterodyne efficiency on the wavefront aberration of a coherent communication system. When the wavefront-aberration root-mean-square (RMS) value was 0.1λ, the heterodyne efficiency decreased about 31%, when compared with an ideal beam with no wavefront aberration. In the same year, Das et al. [11] designed a hexagonal-array coherent detector to address the mismatch between the signal phase and local oscillator phase, which effectively improved the heterodyne efficiency and detection signal-to-noise ratio.

In 2000, Belmonte and Rye [12] analyzed the influence of atmospheric turbulence on the power performance of coherent heterodyne light by using numerical-simulation technology. When the transmission distance increased, the main influence brought by atmospheric turbulence was beam diffusion, which led to a serious decline in coherent received power.

In 2003, Belmonte [13] analyzed the signal degradation in a heterodyne-detection receiver caused by atmospheric turbulence. When a laser beam with a short

wavelength was transmitted through medium-strong turbulence, the smaller receiving aperture was basically the same as the larger receiving aperture, in terms of reducing the degree of coherent diffraction loss.

In 2004, Taylor [14] proposed a coherent-detection method using digital signal-processing technology, and experimentally demonstrated digital coherent-light detection using differential phase-shift keying (DPSK) modulation with a transmission rate of 10 Gbit/s. This method achieved the same high-sensitivity detection as zero-difference detection.

In 2006, Lange et al. [15] tested a binary phase-shift keying (BPSK) modulated communication link with 142-km zero-difference detection. The zero-difference detection link was free from background noise, and the transmission rate reached 5.625 Gbit/s.

In 2006, Toyoshima [16] studied the optimal relationship between a focused beam and the optical-fiber mode-field size in the presence of random angular jitter. When the ratio of the normalized random angular jitter to the mode-field radius was greater than 0.3, the average bit-error rate (BER) of the system was significantly reduced.

In 2008, Belmonte et al. [17] studied the influence of scintillation and phase distortion caused by atmospheric turbulence on the performance of optical wireless communication links. The study showed that high-order phase compensation could effectively improve the sensitivity of the receiving end. In the same year, Tsiftsis [18] derived the analytic expression of the interrupt probability and mean bit-error rate of a heterodyne-detection system under atmospheric turbulence. The results showed that the bit-error rate of the heterodyne-detection system could not reach 10^{-9} in practice.

In 2009, Sandalidis et al. [19] derived the bit-error rate expression under atmospheric-turbulence conditions, and further studied the bit-error performance of a optical wireless coherent communication system using differential phase-shift keying under different turbulence intensities. Numerical simulation proved that the bit-error rate of the system gradually increased with the increase of turbulence intensity.

In 2010, Djerroud et al. [20] conducted an experiment on a 5-km atmospheric-turbulence radio optical-coherent communication link and found that the phase noise of the communication link was mainly caused by atmospheric turbulence and mechanical jitter. In the same year, Salem et al. [21] derived an analytical expression for the heterodyne efficiency after frequency mixing using partially coherent Gauss–Schell model beams, and discussed the influence of wavefront deactivation and the detector radius on the heterodyne efficiency. They found that partially coherent beams were less affected by wavefront deactivation than coherent beams.

In 2011, Fice et al. [22] analyzed and evaluated in detail the Fice phase-locked synchronous homodyne coherent receiver. The evaluation results showed that a synchronous homodyne coherent receiver with a local vibro-optical phase lock could effectively recover the carrier of coherent optical detection.

In 2012, Bayaki et al. [23] established a optical wireless communication system model based on coherent and differential space-time coding, and compared it with an

intensity modulation/direct-detection (IM/DD) optical wireless communication system model. They found that the optical wireless communication system based on coherent and differential space-time coding obtained greater performance gains.

In 2016, Kikuchi [24] studied the application of digital signal-processing technology in a coherent receiver and analyzed the quantum-noise characteristics of its coherent detection. They found that the combination of coherent detection and digital signal processing provided effective compensation for the carrier recovery of a coherent receiver.

In 2022, Kakarla [25] studied receiver sensitivity using multipulse position modulation (M-PPM) and designed a communication link for a phase-sensitive optical-amplification receiver using M-PPM combined with quadrature phase-shift keying (QPSK) modulation. The results showed that when the BER $= 10^{-3}$, the optimal sensitivity of 2.1 dB could be provided by using the modulation mode combined with 64-PPM and QPSK.

In addition, the above literature shows that coherent detection [3] has very strict requirements on the detection mode [9, 14, 15, 23–25], laser light source [4, 8, 21], channel and other environmental factors [6, 12, 13, 18–20], and performance of the detector at the receiving end [11, 22]. In coherent detection, a suitable laser light source is the key to detection. Secondly, to ensure the optimal coherence detection sensitivity, polarization matching [11] and wavefront matching [10, 17] should be performed on the laser beam at the receiving end. The actual beam propagation and drift caused by atmospheric turbulence [16] also affects the efficiency of coherence detection, thus limiting the application of optical wireless coherence communication technology.

Domestic Research Progress

In the 1980s, with the rise of research in the field of optical wireless coherence communication, domestic scholars conducted in-depth research on coherence-detection technology. The optical-field distribution, atmospheric turbulence, polarization control, and phase matching all have significant effects on coherence detection, and scholars conducted in-depth theoretical research in this direction.

In 1985, Sun [26] proposed a 10.6-μm laser heterodyne receiver scheme, and experimentally tested the heterodyne-detection efficiency under different turbulence intensities. When the aperture of the receiving system D and atmospheric coherent length equaled r_0, the turbulence caused by the signal-power ups and downs was minimal and the heterodyne efficiency was the largest.

In 1988, Qian [27] analyzed the analytical relationship between the local oscillator gain, sensitivity, and power, and discussed several limiting factors of the local oscillator power. He found that the greater the local oscillator power, the greater the sensitivity improvement, and the closer the signal-to-noise ratio was to the quantum detection limit. The coherent optical heterodyne detection improved the receiving sensitivity by 10–30 dB, compared with the IM/DD mode.

In the same year, Zhang and Lin [28] conducted heterodyne-detection experiments and found that when a photodiode was used for heterodyne receiving, the optimal local-vibration power was related to the size of the applied bias. That is, using a higher bias in the operating-bias range could effectively improve the local-vibration power and avoid coherent-detection efficiency saturation.

In 1992, Zhang et al. [29] carried out a detailed analysis on the application of a limit-frequency discriminator (L-D) in a coherent receiver based on M-ary continuous-phase frequency-shift keying (CPFSK). The optimal modulation index was one when the beat linewidth of the normalized intermediate-frequency (IF) signal was no less than 2%.

In 1996, Min et al. [30] designed a polarization controller composed of four fiber-optic compounders to meet the polarization stability requirements in coherent communication. The fluctuation range of the output optical power was less than 1%, which met the requirements of practical communication.

In 1999, He [31] designed a coherent Lidar detection system based on the scattering-echo characteristics of targets and the phase-matching principle. The system successfully received the target echo signal 7.8 km away, after adding an adjustable beam expander in the local vibration path.

In 2003, Wang et al. [32] established a heterodyne-detection system for Gaussian and Airy speckle laser distribution, and analyzed the influence of wavefront loss on the differential efficiency. The simulation results showed that the maximum heterodyne efficiency between an Airy spot and a Gaussian beam was 0.82.

In the same year, Xiao et al. [33] analyzed the practical application of coherent-detection technology in an optical-fiber communication system. When a 1550-nm laser beam was used for transmission, the coherent-detection technology improved the sensitivity of the receiving end by 20–23 dB, and the transmission distance reached more than 100 km.

In 2005, Li et al. [34, 35] designed polarization controllers based on a simulated-annealing algorithm (SA) and a genetic algorithm (GA) to control the stability of the polarization state at the receiving end. For the polarization controller with the SA algorithm for feedback control, the output optical-power fluctuation range was less than 1%. After the GA feedback control, the output optical power fluctuation was less than 2%.

In 2007, Li et al. [36] studied the impact of arrival-angle fluctuation on the interference efficiency of the differential detection system, and experimentally tested the interference efficiency and signal-to-noise ratio of the system. Experiments showed that, under the influence of atmospheric turbulence, the beam arrival angle fluctuated obviously, and the signal-to-noise ratio decreased by 20 dB.

In the same year, Xu et al. [37] studied the spatial-coherence degradation caused by atmospheric turbulence and investigated the heterodyne-detection efficiency and signal-to-noise ratio under different turbulence intensities. They found that with the increase of atmospheric-turbulence intensity, the light spots shifted significantly and expanded to different degrees, and the heterodyne-detection efficiency and signal-to-noise ratio showed a decreasing trend.

In 2008, Zhang et al. [38] conducted a heterodyne-detection experiment and measured the echo signal of the target. When the local optical power remained unchanged and the signal optical power was reduced to a six orders of magnitude difference with the local optical power, the heterodyne signal could still be detected.

In 2010, Yazhou et al. [39] studied the influence of the dynamic tracking error on the mixing efficiency and conducted a numerical-simulation analysis. The simulation analysis showed that the receiver-aperture design of the optical wireless coherent communication system must match the alignment error to obtain the optimal bit-error performance of the link.

In the same year, Gao et al. [40] derived the mathematical model and SNR expression of a balanced detection process, and simulated the model. When the beam-splitting ratio range of the fiber beam splicer at the receiving end was [0.293, 0.707], the SNR of the balanced detection was higher than that of ordinary heterodyne detection.

In 2011, Liu et al. [41] deduced the analytical formula of heterodyne-detection efficiency under different distributions of signal light and local oscillator light for a coherent-communication receiving system. The results showed that the efficiency of heterodyne detection using two beams with a uniform amplitude distribution was about 72%, and the efficiency of heterodyne detection using two beams with Gaussian and uniform distributions was 82%. This further proved the rationality of replacing Gaussian beams with a uniform amplitude distribution. In the same year, Xiaoxi et al. [42] designed a heterodyne receiver using asynchronous demodulation, and experimentally compared it with a system using IM/DD to prove the feasibility of the design.

In 2013, Zhang [43] analyzed the influence of a collimation mismatch on the mixing efficiency of a optical wireless coherent communication system. The research results showed that a collimation mismatch could lead to a serious decline in mixing efficiency, and it was necessary to minimize the increase of the loss supporting role between the signal light and local light.

In the same year, Wu et al. [44] studied the influence of wavefront distortion on the signal-to-noise ratio of a BPSK modulated coherent-communication link in zero-difference detection under different atmospheric-turbulence conditions. The experimental results showed that the phase distortion caused by atmospheric turbulence destroyed the phase matching of the beam at the receiving end, resulting in a sharp decline in the mixing efficiency and a decrease in the signal-to-noise ratio of the system.

In 2014, Li et al. [45] studied the influence of light-source parameters on the beam-transmission polarization characteristics under atmospheric turbulence. When the transmission distance was short, the beam polarization was mainly affected by the light-source parameters. When the transmission distance was longer, the beam polarization was obviously affected by atmospheric turbulence.

In the same year, Liu et al. [46] designed and tested a balanced detector using a photodiode and transresistance amplifier. Experimental results showed that the balanced detector improved the signal-to-noise ratio of the system by about 10 dB in the coherent communication link.

In 2015, Li et al. [47] studied the heterodyne-detection efficiency of a Gaussian light distribution and conducted numerical simulations. They found that the heterodyne-detection efficiency was the best when the detector surface radius was greater than twice the waist radius of the local oscillator beam, and the waist radius of the signal beam was ensured to equal the local oscillator beam.

In the same year, Dai et al. [48] analyzed the formation mechanism of channel noise and designed 2.5-Gbit/s and 5-Gbit/s balanced-detection detectors that suppressed atmospheric channel noise. The sensitivity of the former coherent homodyne detection reached −41.2 dBm, and the sensitivity of the latter reached −34.3 dBm.

In 2016, Chang et al. [49] analyzed the working mechanism of a covariance coherent receiver using a Costas phase-locked loop and established a mathematical model for it. When the transmission rate was 5 Gbit/s, the receiving sensitivity reached −41.4 dBm.

In the same year, Wang et al. [50] studied the distance-measurement theory of the heterodyne principle of a double optical-frequency comb, and numerically simulated the relationship between the frequency-comb repetition-frequency difference and the ranging performance. When the signal jacks were sampled in time, the smaller the repetition-frequency difference was, the greater the cumulative noise was, and the worse the ranging accuracy was.

In 2017, Ma et al. [51] studied the relationship between the polarization state of the signal light and the mixing efficiency in coherent detection, and conducted numerical simulations. The maximum mixing efficiency could be obtained when the received signal light was in a circular polarization state.

In the same year, Nan et al. [52] studied the influence of spot-size deviation and optical-axis offset caused by optical elements on the mixing efficiency of the receiving end. It was found that when the spot-size deviation in the spatial optical mixer was less than 10 μm, it usually had little effect on the mixing efficiency. However, when the optical-axis offset range caused by the optical element in the spatial optical mixer was greater than 100 μrad, the effect on the mixing efficiency was usually greater than the spot-size deviation.

In 2018, Lei et al. [53] established an optical wireless communication system model based on the principle of coherent detection, and proved through numerical simulation that the receiver sensitivity of the optical wireless communication system could reach −49 dBm, which had high practical-application value.

In 2019, Feng et al. [54] studied the influence of light sources with different linewidths on the coherent-detection performance, and conducted field experiments on laser light sources with three linewidths. When conducting medium- and long-distance detection, the detection effect of light sources with linewidths of dozens of kilohertz was good.

In 2021, Sun et al. [55] established a random attenuation model of a satellite–ground laser-link channel and analyzed the bit-error rate performance of the model in heterodyne detection. The results showed that the link zenith angle should be less than 20° as far as possible to ensure good link communication.

Researchers in China have investigated coherence detection in terms of the optical-field distribution [32, 41, 47, 51], atmospheric turbulence [26, 36, 37, 44, 54, 55], polarization control [30, 34, 35, 45], phase matching [31, 43], and detector performance at the receiving end [27–29, 40, 42, 46, 48, 49, 53]. Compared with direct detection, the sensitivity and signal transmission capacity of coherent detection have significantly improved since the development of coherent-detection technology.

In domestic coherent-detection technology research, Xi'an University of Technology has also made good achievements. The polarization controller designed by Min [30], Li [34], and others stabilized the output optical power within 1%, effectively ensuring the polarization stability and improving the coherence-detection sensitivity.

Wang [32] and Liu [41] et al. summarized the different influences of the light-source distribution on the coherence detection efficiency. Li [47] and Feng [54] et al. conducted an in-depth study on how the linewidth of the light source affected the dry-detection performance of the receiving terminal. Xiao [33] demonstrated that the transmission distance of a coherent communication system using coherent detection could reach more than 100 km.

Experiment on Coherent Detection in Radio Optical Communication

Optical wireless coherent communication systems use a coherent-detection mode at the receiving end to obtain the best mixing efficiency and system signal-to-noise ratio. This helps to improve the detection sensitivity of coherent communication systems and detect weak signals.

In recent years, the research group of Professor Ke Xizheng from Xi'an University of Technology has extensively researched the influence of optical wireless coherent communication system performance on the optical-communication system design, signal transmission-code modulation, received-signal detection mode, and other aspects, and has made remarkable achievements. A space coherent optical communication system terminal was successfully developed, and an experiment on a optical wireless coherent communication system with a 1.3-km transmission distance was completed.

Coherent-Detection Techniques

Coherent-detection technology can improve the receiver sensitivity. Table 21.1 compares several coherent-detection technologies in terms of the difficulty of the engineering implementation and the heterodyne implementation method.

The principle of direct-detection technology is simple. The photodetector directly converts the optical signal detected by the receiving end into an electrical signal. It is suitable for communication systems with low equipment requirements and short

Table 21.1 Comparison of different probe modes [56]

Detection mode	Heterodyne gain	Applicable communication system	Laser wavelength tunability requirements	Polarization matching	Temperature control requirement	System equipment requirements
Nonoptical extraterri-toriality	Gain is generated by adding local light	Subcarrier modulation (SCM)	Does not require	Does not require	Relatively loose	Relatively loose
Coherent light detection	Gain is generated by adding local light	Amplitude-shift keying (ASK), phase-shift keying (PSK), frequency-shift keying (FSK), etc.	Requires a wide tuning range and a frequency discrimin-ator	Requires strict polarization matching	Very strict	Very strict
Self-zero detection	Gain is generated by adding a pre-fiber amplifier	DPSK	Does not require	Does not require	Relatively loose	Very strict
Direct detection	No gain	On-off keying (OOK), SCM	Does not require	Does not require	Relatively loose	Relatively loose

transmission distances, and the receiving sensitivity is low [57]. Coherent-detection technology requires high-precision optical devices and can meet the needs of communication systems with long transmission distances.

Autozero-detection technology has obvious advantages in receiving sensitivity and transmission distance; however, it requires the signal light to strictly match the local light, which is suitable for communication systems with strict limits on the laser linewidth. Nonoptical outfield detection technology overcomes the phase-matching problem, makes up for the shortcomings of direct-detection and coherent-detection technology, and has loose requirements for the communication system equipment.

Analysis of Coherent Detection Systems Under Atmospheric Turbulence

Beam drift, beam scintillation, and arrival-angle fluctuation caused by atmospheric turbulence affect the detection efficiency of coherent-detection systems [58]. Among them, the effect of atmospheric turbulence on light waves can be summarized as amplitude and phase disturbances.

Effect of Arrival-Angle Fluctuation on Coherent-Detection Systems

The arrival-angle fluctuation is the main factor that reduces the optical coherence of a signal. At the receiving end, the mixing efficiency of the signal light subjected to turbulence and the local oscillator light is reduced, and the heterodyne-detection efficiency is affected [59]. The parameters affecting the arrival-angle fluctuation include the zenith angle, turbulence intensity, internal and external atmospheric-turbulence scales, receiving aperture, and the ratio of the receiving-antenna aperture radius to the local beam-waist radius (Ω).

Figure 21.1 shows the effect of the zenith angle on the average mixing efficiency of a coherent-detection system. With the increase in zenith angle, the power of the optical signal received by the receiving end gradually decreases, and the heterodyne-detection efficiency obviously decreases. The larger the zenith angle is, the faster the average mixing efficiency decreases with the transmission distance. Therefore, the larger the zenith angle is, the greater the impact of turbulence will be, resulting in a more obvious decline in heterodyne-detection efficiency [60].

The results of this study are the same as those given in [55]. The researchers in [55] also carried out numerical simulations on the variation of the bit-error rate of a single-input single-output (SISO) system with the zenith angle. The results showed that the bit-error rate of the system increased with the increase of the zenith angle.

Figure 21.2 shows the influence of the turbulence intensity on the average mixing efficiency of a differential detection system. When the atmospheric refractive-index structure constant differs, the change of turbulence intensity significantly decreases the average mixing efficiency caused by the fluctuation of the arrival angle, and the heterodyne-detection efficiency also decreases [60].

Figure 21.3 shows the influence of the internal and external turbulence scales on the average mixing efficiency of an atmospheric-turbulence detection system. Under

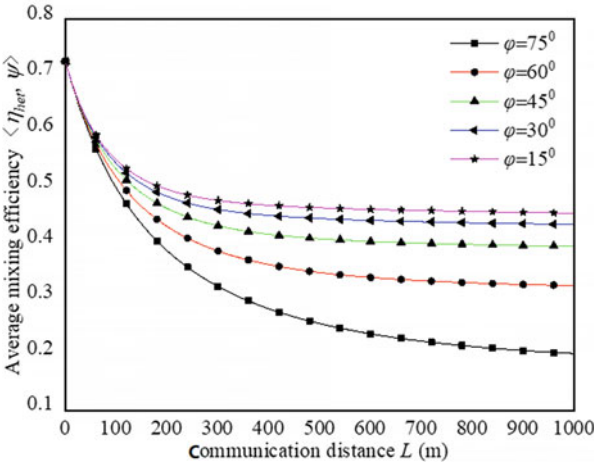


Fig. 21.1 Average mixing efficiency of different zenith angles varies with the transmission distance [60]

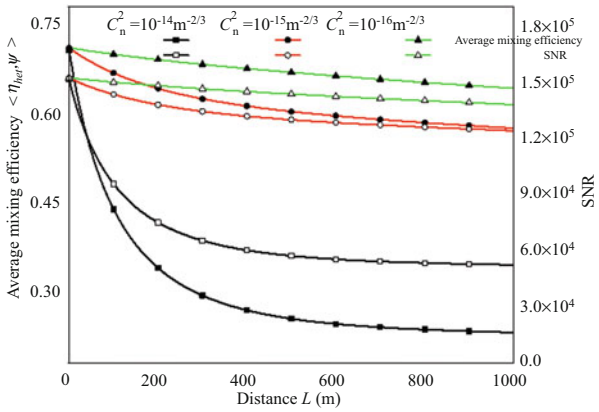


Fig. 21.2 Change of the average mixing efficiency with transmission distance L with different atmospheric refractive-index structure constants [60]

different atmospheric-turbulence intensity conditions, the average mixing efficiency increases with the increase of the internal turbulence scale. As the outer turbulence scale increases, it gradually decreases. This is because the variance of the arrival-angle fluctuation is directly proportional to the outer scale and inversely proportional to the inner scale [60].

Figure 21.4 shows the effect of the receiving aperture and Ω on the mixing efficiency of differential detection. Under medium-strong turbulence conditions, the average mixing efficiency decreases with the increase of the transmission distance, owing to the increase in the receiving aperture. Under weak-turbulence conditions, the gap between different Ω-value curves increases with the increase of receiving aperture.

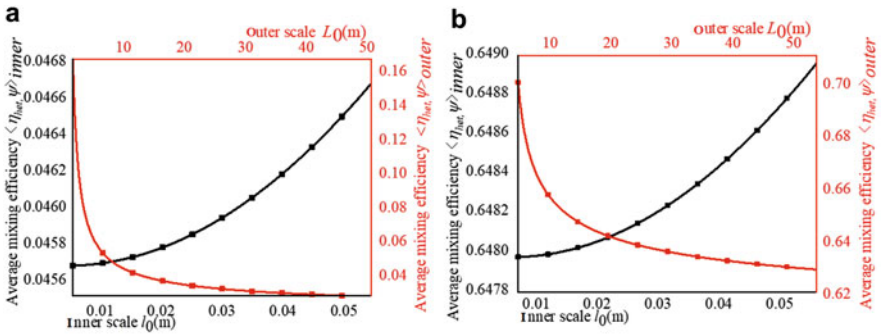


Fig. 21.3 Average mixing efficiency changes with internal/external turbulence scales under different atmospheric-turbulence intensities [60]. (**a**) $C_n^2 = 10^{-14}\ \mathrm{m}^{-2/3}$ (**b**) $C_n^2 = 10^{-16}\ \mathrm{m}^{-2/3}$

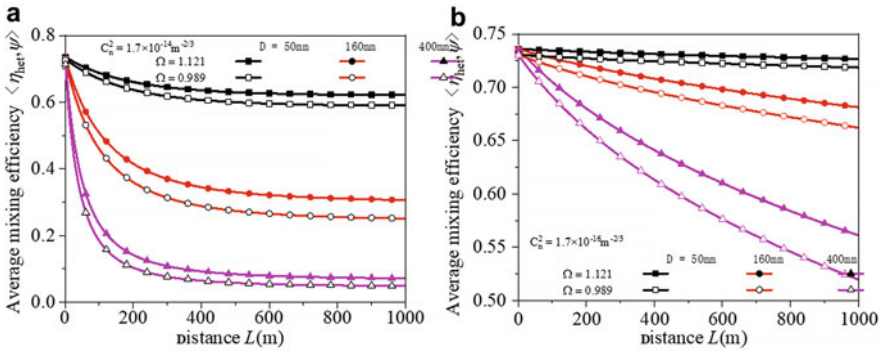


Fig. 21.4 Changes of the average mixing efficiency with transmission distance L under different turbulence conditions, when Ω and the receiving aperture have different values [60]. (**a**) $C_n^2 = 1.7 \times 10^{-14}\ \mathrm{m}^{-2/3}$ (**b**) $C_n^2 = 1.7 \times 10^{-16}\ \mathrm{m}^{-2/3}$

To suppress turbulence effects and improve the heterodyne efficiency of a coherent-detection system, an optical receiving system with an appropriately small diameter should be selected to meet the performance requirements of the communication system when realizing long-distance communication [60].

Effect of Amplitude Fluctuation on Coherent Detection Systems

The amplitude fluctuation affects the performance of an aberration detection system, and the influence is more complicated when phase and amplitude fluctuations coexist. Figure 21.5 shows the relationship diagram of the signal-to-noise ratio of heterodyne detection under the influence of amplitude fluctuation.

The variation of the SNR with the zenith angle is shown in Fig. 21.6, where Fig. 21.6a is the relation curve between the SNR and zenith angle under amplitude disturbance, and Fig. 21.6b is the relation curve between the SNR and zenith angle under amplitude and phase disturbances. The simulation results show that, with the

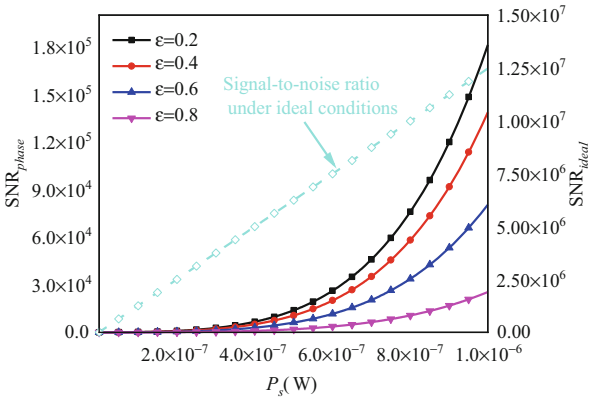


Fig. 21.5 Variation of the SNR with the signal optical power under ideal conditions and different occlusion ratios [61]

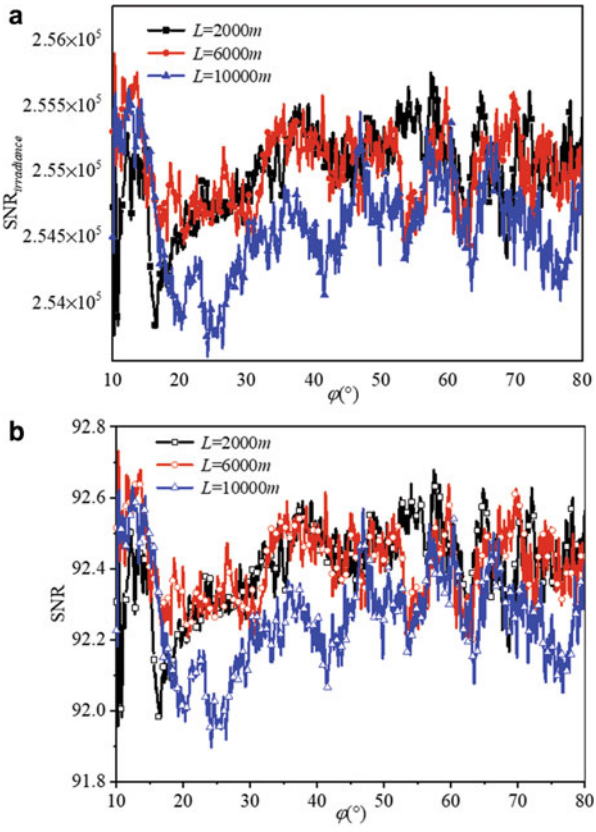


Fig. 21.6 Changes of SNR with the zenith angle under different disturbance conditions and transmission distances [61]. (**a**) Irradiance fluctuation (**b**) Irradiance and phase fluctuation

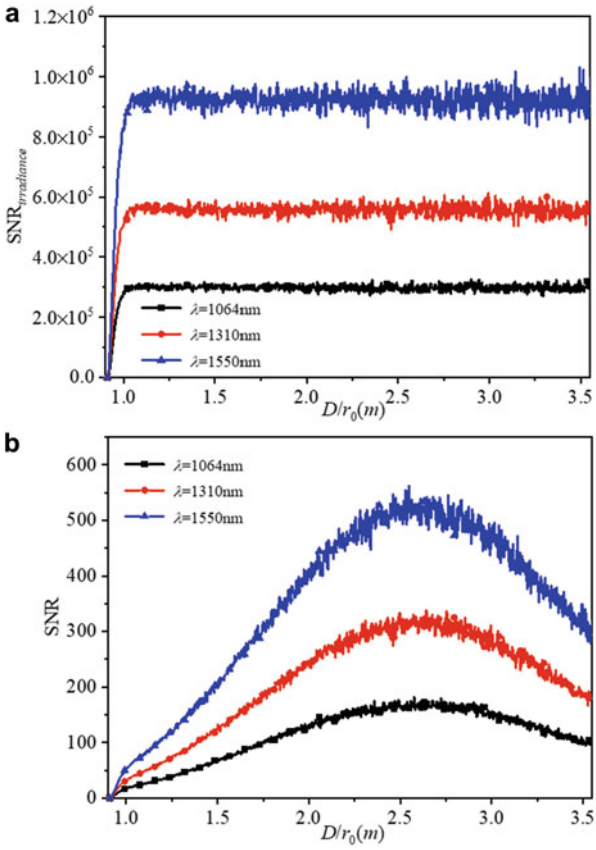


Fig. 21.7 Change of SNR with D/r_0. (**a**) Amplitude perturbation; (**b**) amplitude and phase perturbation [61]

increasing zenith angle, the signal-to-noise ratio of the heterodyne-detection system appears to fluctuate, but the overall trend is decreasing. This is because the distance the signal light travels in relatively strong turbulent regions increases as the zenith angle increases.

The variation of the SNR with atmospheric turbulence D/r_0 is shown in Fig. 21.7, where Figure 21.7a is the relation curve between the SNR and D/r_0 under an amplitude disturbance, and Figure 21.7b shows the relation curve between the SNR and D/r_0 under amplitude and phase disturbances. The simulation results show that the signal-to-noise ratio increases as the wavelength λ increases; therefore, the heterodyne-detection system should choose 1550 nm as the communication frequency band.

1.3-km Heterodyne-Detection System Experiment

The authors of Reference [61] established a 1.3-km heterodyne-detection communication link, and the system block diagram is shown in Fig. 21.8. At the transmitter end, a 1550-nm beam was used as the signal light. The source information was

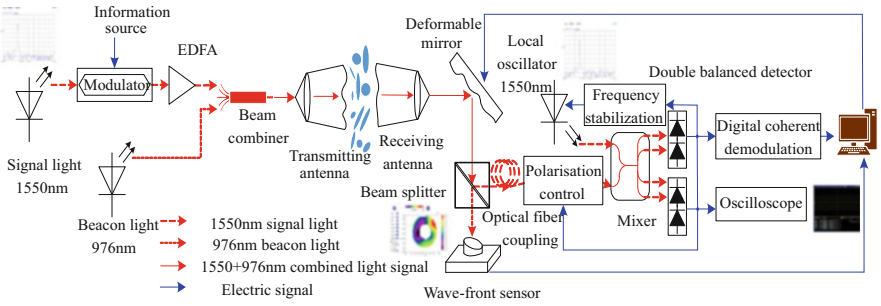


Fig. 21.8 Schematic diagram of the heterodyne-detection system

loaded onto the signal light through the phase modulator and fed into the transmitting optical system through an erbium-doped fiber amplifier (EDFA).

When transmitted in the atmospheric channel (1.3 km long), it was affected by turbulence, resulting in the dispersion of the received light spot and an increase in the system bit-error rate. In severe cases, the communication link might be interrupted [62]. Therefore, at the receiving end of the coherent-detection system, adaptive-optics (AO) technology was needed to correct the optical wavefront distortion of the signal [63].

After the correction, the signal light was mixed with the local oscillator light, and the baseband signal was recovered in real time using polarization control, laser-frequency stabilization, coherent detection, and digital coherent demodulation. The 1.3-km detection experiment was completed.

The experimental results showed that when the polarization control system closed the loop, the amplitude jitter caused by polarization obviously decreased, and the signal quality at the receiving end was effectively improved [61]. Under different weather conditions, the center frequency of the IF signal changed around 82 MHz, which was consistent with the control accuracy of the laser-frequency stabilization module [61]. After the AO system closed the loop, the quality of the IF signal was effectively improved, and the primary and secondary harmonics of the corrected IF signal power spectrum were “sharper” [64].

Figure 21.9 shows the variation of the IF signal amplitude measured by the 1.3-km detection experiment. As shown in Fig. 21.9a, the mixing efficiency under the influence of the polarization control increases by about 64%, and the coherence-detection efficiency is improved. The measured results in Fig. 21.9b are similar to those in Fig. 21.9a. The simulation results show that when the polarization control system closes the loop, the amplitude jitter caused by polarization obviously decreases, and the signal quality of the receiver is effectively improved.

Figure 21.10 shows the power-spectrum variation curve of the IF signal in open and closed loops of the AO system under different weather conditions. An analysis of the influence of heterodyne detection on the IF signals indicates that the AO system can effectively improve the IF signal quality after the closed loop, and the

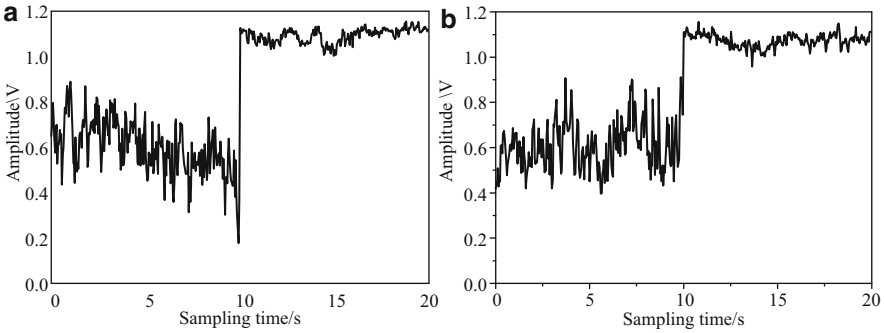


Fig. 21.9 Amplitude changes of the IF signal in open and closed loops of the polarization control system [61]. (**a**) 2018-04-23. (**b**) 2018-04-24

primary and secondary harmonics of the corrected IF signal power spectrum are "sharper" [64].

At the receiving end of the heterodyne-detection system, the IF signal light is mixed with the local oscillator light; the output is shown in Fig. 21.11. Figure 21.11a is the amplitude change curve of the IF signal, and Fig. 21.11b is the change curve of the power-spectrum analysis. The simulation results show that the center frequency of the IF signal changes around 82 MHz, which is consistent with the control precision of the laser-frequency stabilization module.

Analysis of a Coherent-Detection System Under the Influence of Channel Noise

Optical wireless coherent communication has a good effect on long-distance communication; however, it is generally affected by background optical noise, and affects the dry-detection efficiency of the receiving terminal. In addition, considering the influence of dark-current noise and electronic thermal noise [65], we focus on the noise statistical characteristics of the signal and various random circuit noises on the detection system, and estimate the performance.

Signal-to-Noise Ratio Analysis

Figure 21.12 shows the relationship curve between the signal-to-noise ratio and received optical power of the coherent-detection system. With the increase of the received optical power, the signal-to-noise ratios of coherent detection and direct detection under ideal conditions and direct detection affected by noise are compared and analyzed. The simulation results show that the signal-to-noise ratio of coherent detection is 4.1 dB higher than that of direct detection under ideal conditions, and the direct-detection signal-to-noise ratio exceeds nearly 23 dB under specific thermal noise and loose-particle noise.

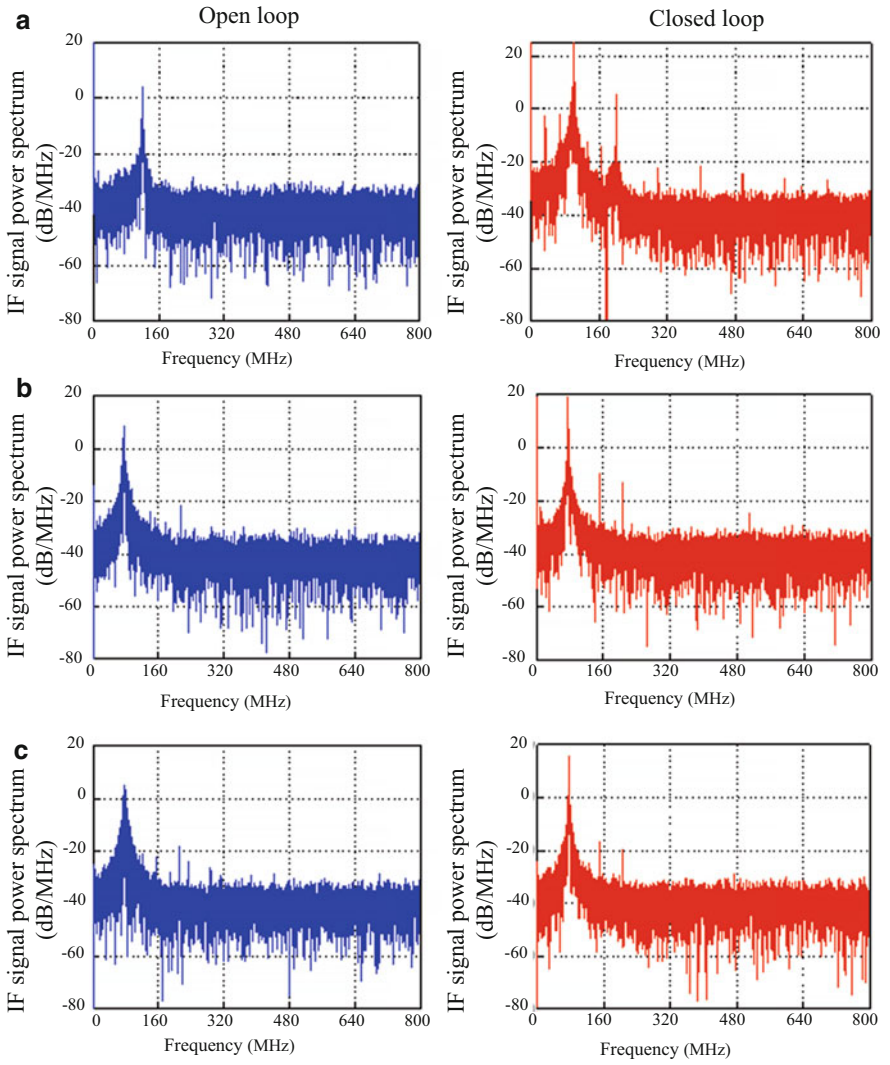


Fig. 21.10 Changes of the IF signal power spectrum in open and closed loops under different weather conditions [61]. (**a**) Cloudy, (**b**) sunny, (**c**) light rain

Analysis of the Influence of Relative-Intensity Noise

The constellation of different local lasers and relative-intensity noise (RIN) is shown in Fig. 21.13. Analysis results show that with the increase of RIN, the received signal-to-noise ratio decreases, the bit-error rate increases, and the coherence-detection efficiency is seriously affected. Under the same RIN influence, the coherence-detection efficiency decreases with the increase in local light power.

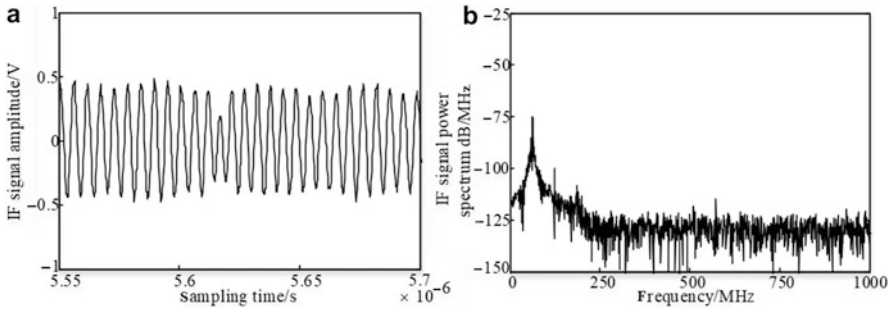


Fig. 21.11 IF signal output [61]. (**a**) The waveform of IF signal. (**b**) The power spectrum of IF signal

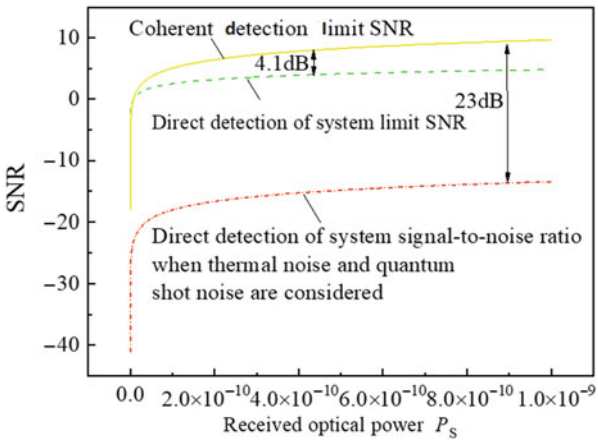


Fig. 21.12 Relationship between the received optical power and system SNR [66]

Thermal Noise–Effect Analysis

Figure 21.14 shows the QPSK demodulation signal constellation with different signal-to-noise ratios, when only thermal noise is considered. Figure 21.14a shows the constellation diagram of the QPSK demodulation signal when the signal-to-noise ratio is 30 dB. The constellation points at this time are easy to identify. Figure 21.14b corresponds to a signal-to-noise ratio of 10 dB.

When the signal-to-noise ratio in Fig. 21.14c is 1.5 dB, the constellation points are relatively dispersed, and the demodulation signal is seriously affected by noise. In Fig. 21.14d, in the case of the −1.5-dB signal-to-noise ratio, the distribution of the constellation points is very fuzzy, indicating that the dry-detection performance of the receiving terminal is seriously damaged. The analysis can be done when there is only thermal noise influence. When the signal-to-noise ratio decreases gradually, the degree of the constellation-map dispersion increases.

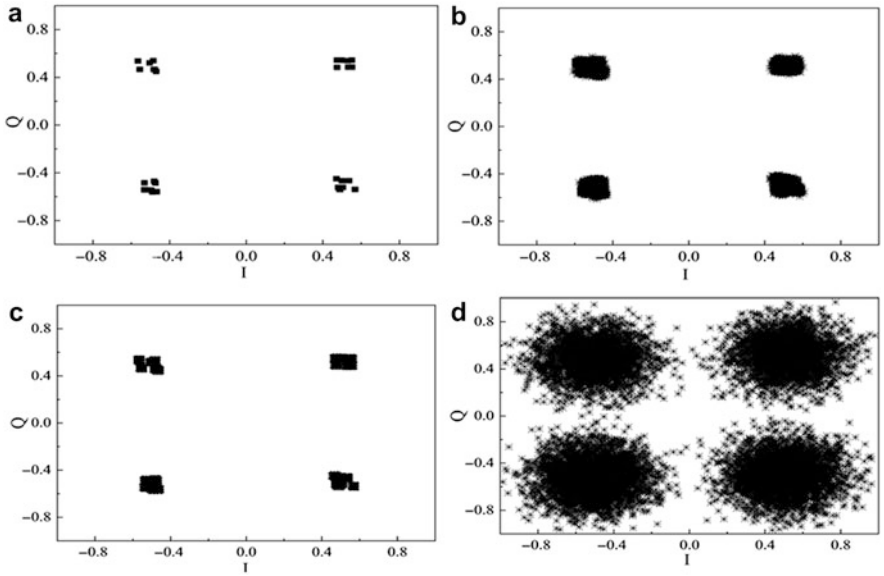


Fig. 21.13 Constellation of laser-power and relative-intensity noise with different local vibrations [66]. (**a**) RIN = −160 dB/Hz, PL = 0 dBm (**b**) RIN = −130 dB/Hz, PL = 0 dBm (**c**) RIN = −160 dB/Hz, PL = 10 dBm (**d**) RIN = −130 dB/Hz, PL = 10 dBm

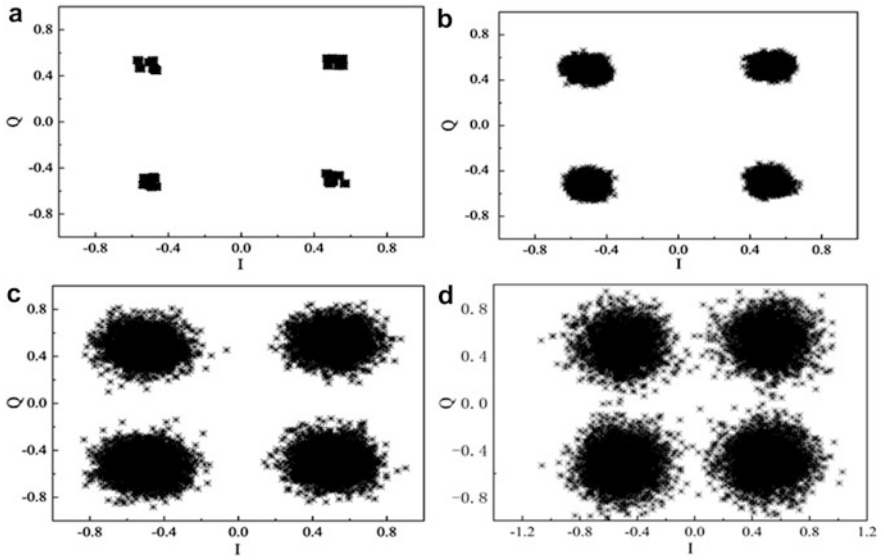


Fig. 21.14 Constellation comparison of QPSK demodulation signals under different SNR [66]. (**a**) SNR = 30 dB (**b**) SNR = 10 dB (**c**) SNR = 1.5 dB (**d**) SNR = −1.5 dB

Analysis of the Influence of Optical-Signal Polarization on Coherent Detection Systems

The polarization of an optical signal has a direct influence on the coherent-detection system, and the polarization control algorithm also has a direct influence on the response time of the polarization control. This section introduces the advantages of the simulated annealing (SA), particle-swarm optimization (PSO), and single-particle optimization (SPO) algorithms, and compares them.

Advantages of the Three Algorithms

Figure 21.15 shows the optimization performance of the SA algorithm corresponding to V_{Pol} for different signal polarization-state change rates. As V_{Pol} decreases gradually, the IF-signal amplitude fluctuates less after convergence. When $V_{Pol} = 0$ rad/s, the amplitude fluctuation of the IF signal after convergence tends to stabilize with the increasing number of iterations, and the coherent-detection system shows good performance.

Figure 21.16 shows the optimization performance of the PSO algorithm corresponding to the polarization-state change rate of different signal lights V_{Pol}. With the increase in the number of iterations, the amplitude fluctuation of the IF signal stabilizes after convergence. Even if the polarization state of the signal light fluctuates, the PSO algorithm has better optimization performance.

The SPO algorithm was designed by combining the SA, PSO, and genetic algorithms. As shown in Fig. 21.17, in the face of rapid changes in the signal-light polarization state, the SPO algorithm shows a good control effect and can quickly stabilize the amplitude fluctuations of the intermediate-frequency signal [68].

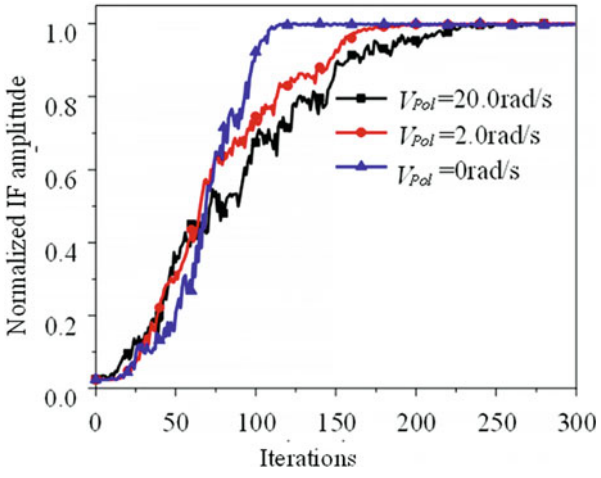


Fig. 21.15 Influence curve of the simulated annealing algorithm on the IF signal amplitude [67]

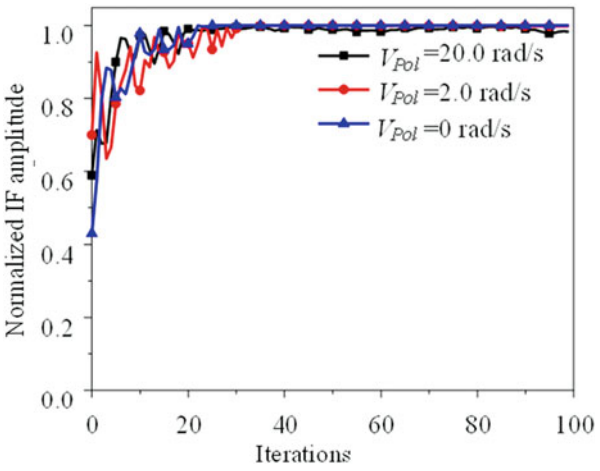


Fig. 21.16 Influence curve of the particle-swarm optimization algorithm on the IF signal amplitude [67]

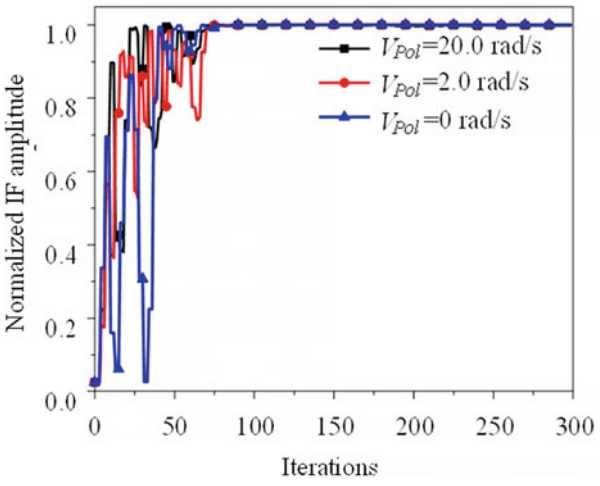


Fig. 21.17 Influence curve of the SPO algorithm on the amplitude of the IF signal [67]

Comparison of the Three Algorithms

From the perspective of the convergence time, the convergence speed of the SPO algorithm is much faster than that of the SA and PSO algorithms. This is mainly because the SA algorithm needs a large number of iterative operations to reach the optimum, and the existence of the second reset of the PSO algorithm makes it consume nearly twice the time in the control process [68].

In terms of the time required to reach the maximum IF signal amplitude, the SA algorithm requires the most convergence time, and the SPO algorithm has obvious advantages in convergence time compared with the others [67]. From the viewpoint

of parameter determination and algorithm complexity, the SPO algorithm is more complex than the SA and PSO algorithms.

Figure 21.18 compares and analyzes the IF-signal amplitude changes with the convergence time of the three algorithms. From the point of view of the time needed to reach the maximum amplitude of the IF signal, the SA algorithm needs the most convergence time. Compared with the SA and PSO algorithms, the SPO algorithm shows obvious advantages in convergence time.

To further verify the advantages of the SPO algorithm, we completed an experimental study of the SPO algorithm in a coherent-detection system. Figure 21.19 shows the experimental results. Figure 21.20a shows the amplitude change of the IF

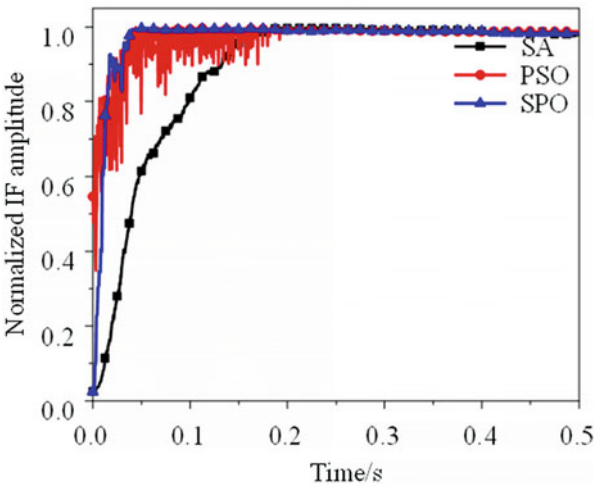


Fig. 21.18 Comparison of the convergence times of the three algorithms [67]

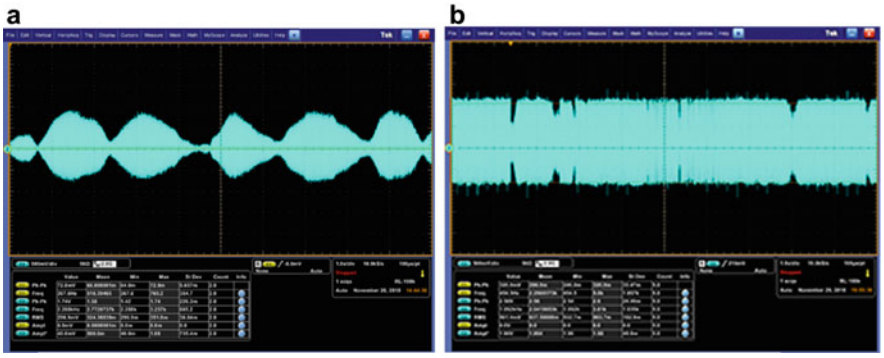


Fig. 21.19 Influence of the polarization control system on the IF signal [67]. (**a**) Open-loop control system. (**b**) Closed-loop control system

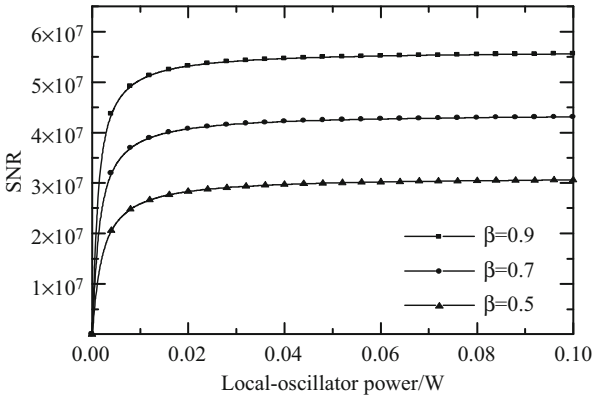


Fig. 21.20 Variation of the signal-to-noise ratio with local optical power [69]

signal output by oscilloscope when the polarization controller PLC-M02 is turned off. It can be seen that the IF signal amplitude changes to a cosine.

Figure 21.20b shows the amplitude change of the IF signal output by oscilloscope when the polarization controller PLC-M02 is turned on. The SPO algorithm is used. It can be seen that the amplitude of the IF signal is basically in a stable state, and the coherence gain at this time is the best.

Performance Analysis of the Light Source on a Coherent Detection System

Analysis of the Local-Vibration Optical Power on the Coherent-Detection System Performance

Figure 21.21 shows the influence curve of the local optical power on the signal-to-noise ratio of the coherent-detection system under ideal conditions. Along with the increase of the degree of response β, the coherent-detection system signal-to-noise ratio increases gradually. With a small change in the local light power, the signal-to-noise ratio of the coherent-detection system increases sharply to the maximum value and is no longer affected by the local light power. The analysis shows that the performance of the coherent-detection system will not change when the local power increases to a specific value.

Figure 21.21 shows the optical power of the coherent-detection system, considering the vibration output signal-to-noise ratio and the influence of the heterodyne current in a nonlinear state, when the parameters of α are not simultaneous [69]. As the α parameters increase gradually, the quantum-noise limited signal-to-noise ratio and heterodyne-detection efficiency decline. As the optical power of the vibration increases, the output current increases gradually, and with the increase of the α parameters, the output current decreases.

Figure 21.22 shows the optical power of the coherent-detection system of the vibration-output signal-to-noise ratio and the influence of the heterodyne current in a

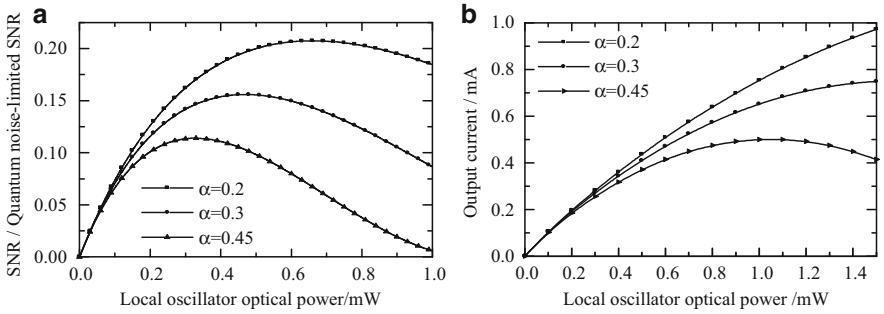


Fig. 21.21 α is not simultaneous when the receiver sensitivity varies with the vibration power relations [69]. (**a**) SNR versus the local oscillator power. (**b**) Output current versus the local oscillator power

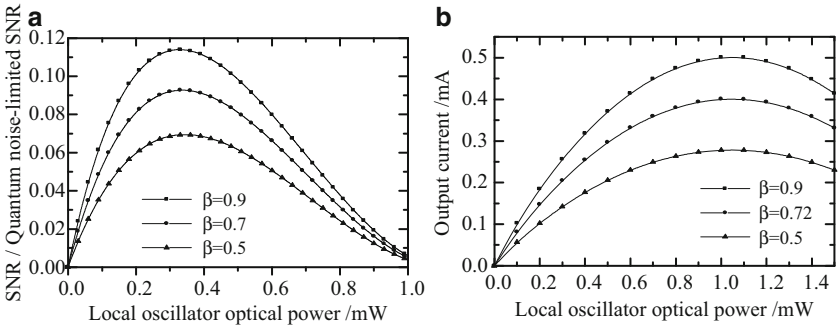


Fig. 21.22 Receiver sensitivity does not vary with the vibration power relations [69]. (**a**) SNR versus the local oscillator power. (**b**) Output current versus the local oscillator power

nonlinear state, when the responsivity β is not simultaneous [69]. With increase in the β response degree, the quantum-noise limit signal-to-noise ratio increases gradually. The output current increases gradually with the increase of the β response degree; in the light of the vibration power of 1 mW, it reached the maximum output current.

Analysis of the Light-Source Linewidth on the Performance of a Coherent Detection System

The small-linewidth drift of a laser will lead to a sharp increase in the transmission bit-error rate of a optical wireless coherent communication link, which has a great influence on the efficiency of coherent detection [70]. Figure 21.23 shows the influences of the light-source linewidth, atmospheric turbulence, transmission distance, and frequency-difference signal bandwidth on the transmission bit-error rate [71].

When the light-source linewidth, atmospheric-turbulence intensity, transmission distance, or bandwidth of the difference-frequency signal increase, the bit-error rate

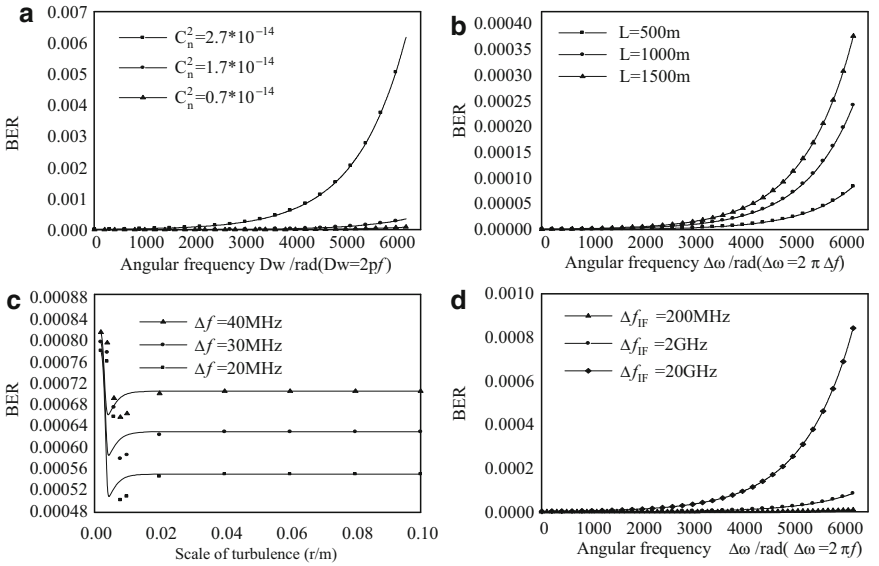


Fig. 21.23 Changing curve of the laser linewidth and BER [71]. (**a**) Laser linewidth effects on BER when turbulence intensity is not the same. (**b**) Laser linewidth effects on BER when transmission distance is not the same. (**c**) Turbulence scale effect on BER when line width is not the same. (**d**) Laser linewidth effects on BER when difference-frequency signal bandwidth is not the same

increases gradually. The simulation results show that the laser linewidth and turbulence intensity have different effects on the coherence-detection efficiency in optical wireless coherent communication systems.

Performance Analysis of a Subcarrier-Modulated (SCM) Nonoptical Outlier-Detection System

In this section, the effects of nonoptical outfield detection (NOD) on the sensitivity of the receiver in ideal and nonideal cases are analyzed.

Ideal-Channel NOHD

In an ideal channel, when an NOHD method using ASK, QPSK, hexadecimal phase-shift keying (16PSK), or other modulation modes is adopted for receiving, the power-change curve of the intermediate-frequency signal is obtained as shown in Fig. 21.24 [56]. These modulation modes all linearly amplify the sent signal, effectively improving the detection sensitivity. As shown in Fig. 21.25, from the signal eye plot after demodulation, when the ASK and QPSK modulation modes are adopted, the BER slope of the signal eye-plot contour is large and sensitive to timing errors. Using 16PSK modulation, the signal eye-map outline is clear. This indicates

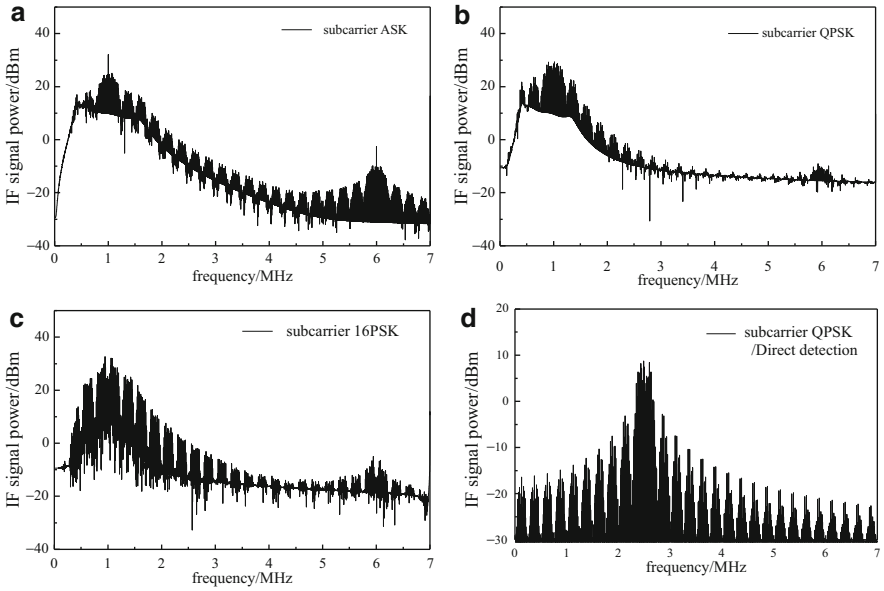


Fig. 21.24 Power spectrum of NOHD received and directly detected signals in an ideal channel with different subcarrier-modulation modes [56]. (**a**) SCM-ASK/NOHD. (**b**) SCM-ASK/NOHD. (**c**) SCM-ASK/NOHD. (**d**) SCM-ASK/NOHD

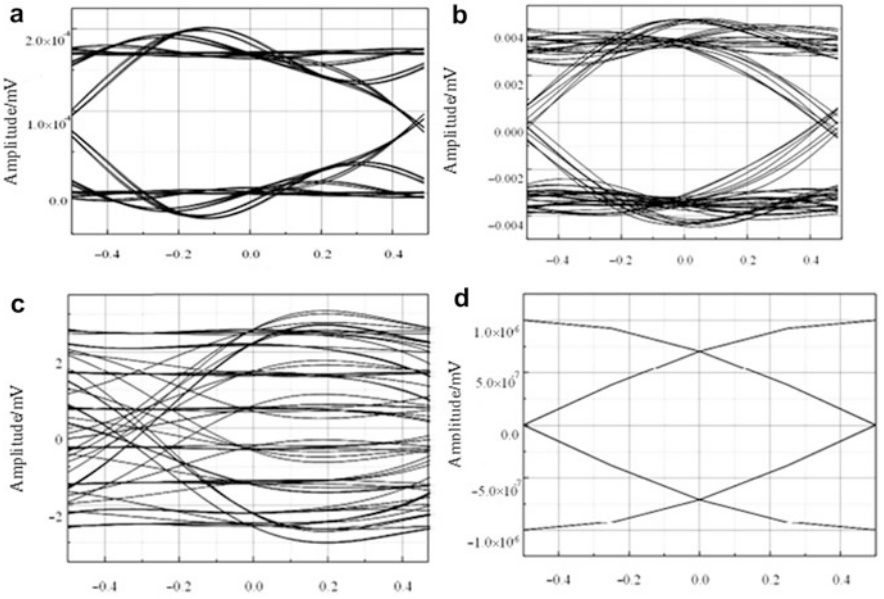


Fig. 21.25 Signal eyes for an ideal channel with different subcarrier-modulation modes under NOHD and direct detection [56]. (**a**) SCM-ASK/NOHD. (**b**) SCM-QPSK/NOHD. (**c**) SCM-16PSK/NOHD. (**d**) SCM-QPSK/DD

that the NOHD mode is suitable for optical-signal reception using various SCM modes.

Nonideal-Channel NOHD

The change curve of the signal-to-noise ratio of the receiver with relative-intensity noise is shown in Fig. 21.26 with Reference [56]. The analysis shows that when the relative-intensity noise is greater than −160 dB/Hz, the detection signal-to-noise ratio shows a decreasing trend, and the detection signal-to-noise ratio curve with a local power of 10 dBm is most seriously affected by the relative-intensity noise. In actual research and analysis, the detection signal-to-noise ratio of the system can be improved by reducing the relative-intensity noise.

Figure 21.27 shows that under the influence of different relative-intensity noise, when the local vibration-light power is greater than a specific value, the receiver's signal-to-noise ratio shows a downward trend, owing to the influence of the relative-intensity noise. The larger the relative-intensity noise, the smaller the signal-to-noise ratio of the receiver.

As shown in Fig. 21.28, when only the thermal noise of the receiver is considered, the signal-to-noise ratio of the receiver increases sharply with the increase in local optical power. When the local optical power is greater than 25 dBm, the signal-to-noise ratio of the receiver tends to stabilize. With the increase in signal optical power, the signal-to-noise ratio of the receiver shows an overall increasing trend, and with the increase of receiver bandwidth, the signal-to-noise ratio gradually decreases.

As shown in Fig. 21.29, when the lognormal distributed-channel model is adopted, in the case of weak turbulence, the received optical power is −30 dBm, the received signal-to-noise ratio is 9 dB, and the bit-error rate is 10^{-5}. This meets the requirements of optical wireless communication. With the increased variance of the light-intensity fluctuation, the average bit-error rate of the system increases gradually and tends to stabilize.

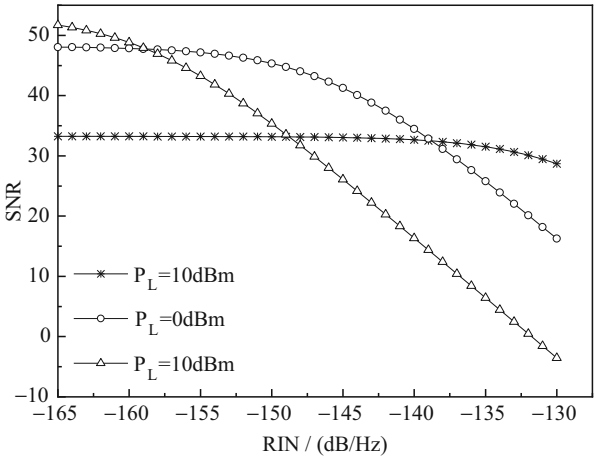


Fig. 21.26 Receiver SNR curve with relative-intensity noise (RIN) value [56]

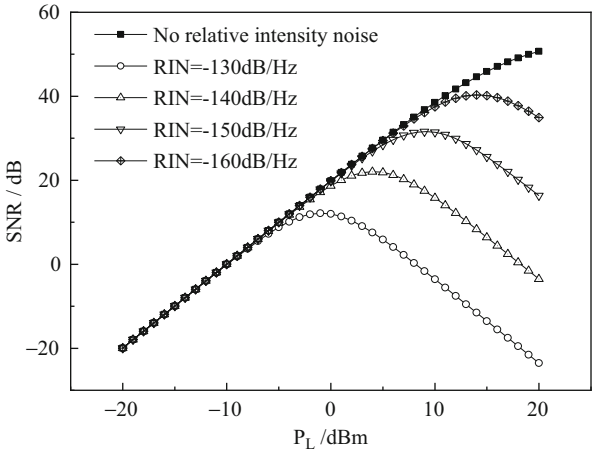


Fig. 21.27 Relation curve between the local optical power and receiver signal-to-noise ratio [56]

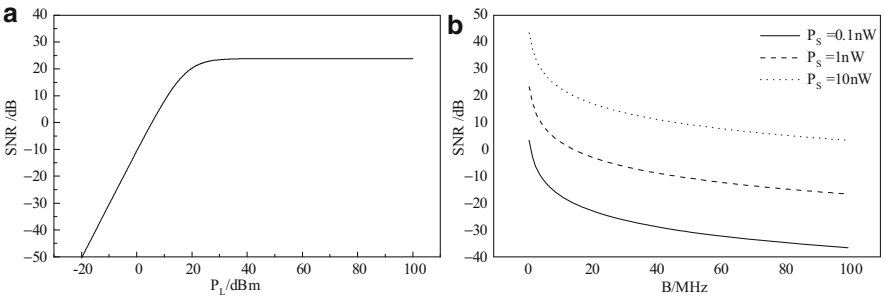


Fig. 21.28 SNR curve of the receiver under the influence of thermal noise [56]. (**a**) Variation of SNR with different values of local oscillator optical power. (**b**) Variation of SNR with different values of receiver bandwidth

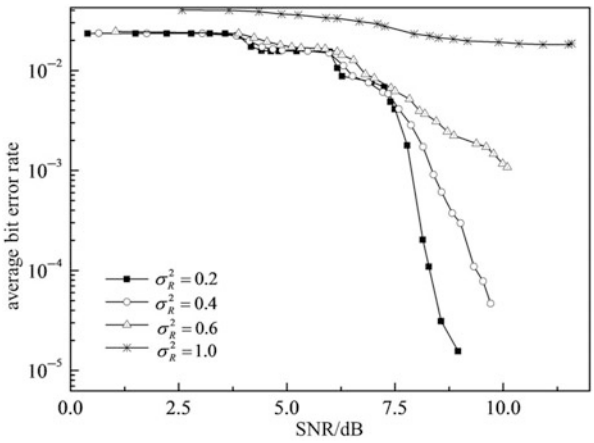


Fig. 21.29 Bit-error rate curve of a lognormal distributed channel under different intensity fluctuation variances [56]

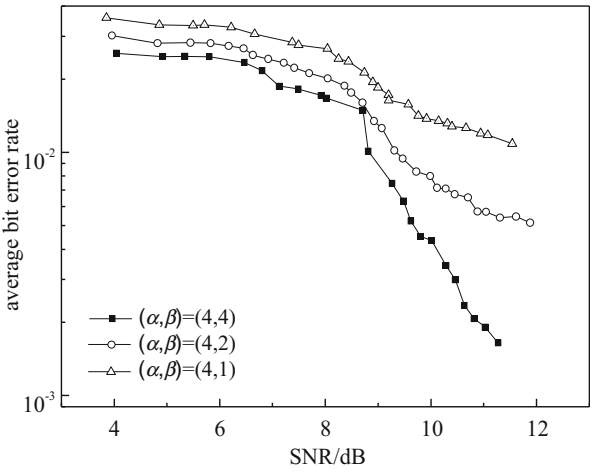


Fig. 21.30 Bit-error rate curves of a gamma–gamma channel with different turbulence intensities [56]

As shown in Fig. 21.30, for the SCM-QPSK/NOHD mode, the bit-error rate of the system increases with the increase of turbulence intensity in a gamma–gamma channel; that is, the phase noise introduced by the atmospheric channel will worsen the signal-to-noise ratio of the system.

NOHD System Experiment

The SCM/NOHD mode has a heterodyne gain approaching 20 dB [72]. As shown in Fig. 21.31, by comparing the NOHD received difference-frequency signal and the direct-detection signal of the SCM orthogonal frequency-division multiplexing (OFDM) signal and their power spectra, it can be seen that the time-domain waveform of the NOHD difference-frequency signal is less affected by noise, and the heterodyne gain of NOHD mode is about 19.2 dB higher than that of direct detection [56].

In the field of optical wireless coherence communication, to detect optical signals and improve the coherence gain, Professor Ke Xizheng's team at Xi'an University of Technology has conducted systematic analyses in the field of optical wireless coherence communication system detection.

1. Kong et al. [73] analyzed the effects of different optical-field distributions on the external difference efficiency under the influence of atmospheric turbulence. They proved that an Airy–Gauss distribution of a signal light and local-vibration light have the best turbulence-suppression effect.
2. Chen [56] proposed a new nonoptical outlier detection method, built an NOHD experimental system, and completed a 1-km communication experiment, proving that the system could obtain a heterodyne gain 20 dB higher than that of direct detection.

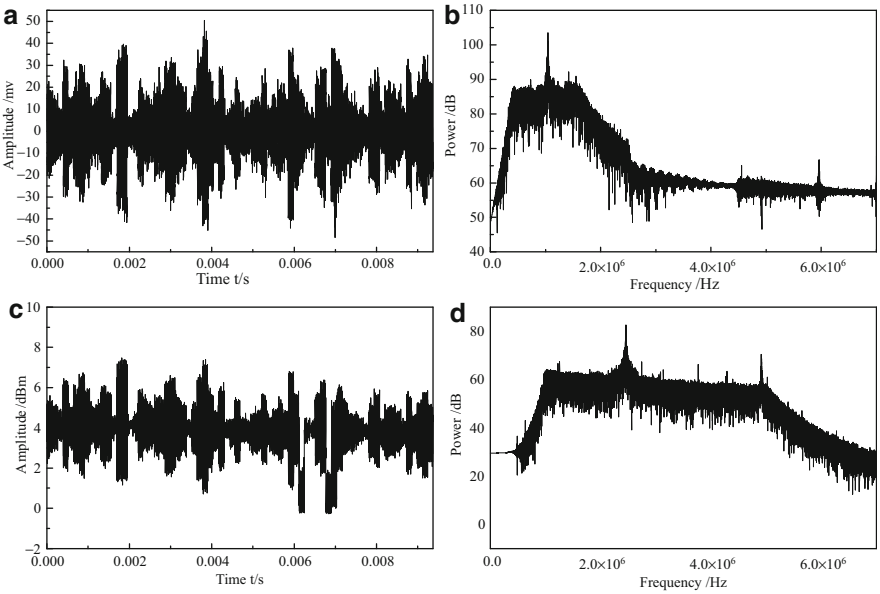


Fig. 21.31 Comparison of the subcarrier difference-frequency signal with the direct-detection signal and power spectrum [56]. (**a**) SCF-OFDM/NOHD, time domain waveform (**b**) SCF-OFDM/NOHD, power spectrum (**c**) SCF-OFDM/DD, time domain waveform. (**d**) SCF-OFDM/DD, power spectrum

3. Yang et al. [74] designed a system using a feedback control-loop circuit. The circuit made the carrier-frequency drift stability within ±2 MHz, and helped the receiver demodulation, which could tune the laser frequency to the laser-frequency tracking signal.
4. Chen [66] studied the change in receiving sensitivity of optical wireless coherent communication under the influence of turbulence, and built a new coherent-detection system. The system's gain was 20 dB higher than that of direct detection.
5. Wu [75] developed a wavefront sense-free heterodyne-detection communication system based on a stochastic parallel-gradient descent algorithm (SPGD). It improved the single-mode fiber coupling power of a optical wireless coherent communication link by about 25 dB.
6. Mao [76] studied the transmission characteristics of a beam in a optical wireless coherent communication system, and analyzed the polarization characteristics of the beam at the receiver, providing a theoretical basis for improving the gain at the receiver.
7. Tan [61] conducted a 1.3-km field experiment. When the IF signal voltage was greater than 260 mV, the bit-error rate was less than 10^{-6}, which effectively completed real-time communication.

8. Wang [77] designed a optical wireless coherence communication demodulation system based on heterodyne detection, and conducted video-transmission experiments with error rates less than 10^{-6} at 400 m and 1.3 km.
9. Ma [67] designed a polarization control scheme using intermediate-frequency signals as feedback signals, and proposed a polarization-control algorithm suitable for optical wireless coherent communication. Experimental results showed that the scheme improved the mixing efficiency by 64%.
10. Bai [78] set up an experimental platform for optical wireless coherence communication, and completed modulation and demodulation experiments of octal phase-shift keying (OPSK) signals at a bit-error rate of 10^{-3} by decoding QPSK intermediate-frequency signals. For details, see Table 21.2.

Color Difference of a Dual-Wavelength Free-Space Optical Communication System

In a free-space optical (FSO) communication system affected by a randomly fluctuating atmospheric refractive index, optical turbulence effects, such as phase distortion, light-intensity fluctuation, beam drift, and arrival-angle fluctuation will occur after a signal beam is transmitted through a free channel [79–81]. Based on real-time detection and correction, adaptive-optics technology can effectively suppress the influence of atmospheric turbulence on a signal light [82].

When the selected signal light has poor transmissibility or severe attenuation in an atmospheric channel, the system transmits a beacon light of other wavelengths through the common optical path with the signal light and detects the wavefront distortion of the beacon light to correct the wavefront distortion of the signal light [83]. However, when Gaussian beams of different wavelengths transmit through atmospheric channels, the angle of refraction and optical-path difference caused by turbulence are different, resulting in a wavefront-distortion difference between the signal light and beacon light, that is, a system color difference [84].

In recent years, extensive research has been conducted on the color difference in dual-wavelength free-space optical communication systems. From the beam-transmission point of view, when a Gaussian beam is transmitted in an atmospheric channel, it will be modulated by atmospheric turbulence and change the propagation direction of the wave vector, resulting in a random change in the beam-propagation direction. In the spatial-frequency domain, the randomly changed direction of the wave vector is constantly superimposed, causing the centroids of the beams in adjacent sampling periods to be in different positions, that is, arrival-angle fluctuation [85].

Italo established a scintillation model of an arrival angle under a non-Kolmogorov turbulence spectrum [86]. Voelz conducted a numerical analysis of the arrival angle of a light beam in strong and weak turbulence regions through ray tracing and wave optics [87]. Conan and Borgnino systematically analyzed the beam angle-of-arrival method on the receiving plane through a theoretical derivation and

Table 21.2 Research work in the field of coherent communication at Xi'an University of Technology

Researcher	Research contents	Research findings
Kong	The influence of atmospheric turbulence on heterodyne efficiency under different light-field distributions was studied	Proved that the Airy–Gauss distribution model had the best antiturbulence effect
Chen	A nonoptical heterodyne-detection method based on subcarrier modulation was proposed	The heterodyne gain of this detection method was about 20 dB higher than that of direct detection
Yang	To solve the problem of carrier-frequency drift, a feedback decision circuit was designed	This circuit stabilized the carrier-frequency drift within ± 2 MHz, which was helpful for demodulating at the receiver
Chen	A optical wireless coherent communication model was established and the influence of channel noise was analyzed	The simulation proved that the heterodyne-detection gain was 20 dB higher than the direct-detection gain
Wu	A wavefront sensorless heterodyne-detection system based on SPGD was designed	The system improved the coupling power of the single-mode optical fiber of a coherent communication link by about 25 dB
Mao	The beam-mode degeneration and polarization characteristics in coherent optical communication were studied	In an oblique-transmission downlink, the beam mode was least affected by turbulence
Tan	The influencing factors of heterodyne-detection performance in optical wireless coherent communication were studied	A 1.3-km video-transmission experiment was completed using the developed heterodyne-detection system
Wang	A demodulation system for radio-optical coherence communication based on heterodyne detection was designed	A video-transmission communication experiment was conducted with a bit-error rate of less than 10^{-6}
Ma	The polarization to frequency-mixing efficiency and a polarization-control algorithm were studied	A small step-backward reset algorithm was designed for polarization control in optical wireless coherent communication
Bai	The demodulation performance of an OPSK modulated optical wireless coherent communication system was verified	Modulation and demodulation experiments of an OPSK signal at a bit-error rate of 10^{-3} were completed

numerical calculation, and explained the spatiotemporal correlation and anisotropic variance of the angle-of-arrival fluctuation [88].

Akira described the correlation between dual-wavelength beams in a weak turbulent region through the cross-correlation power spectrum. Ben-Yosef experimentally measured the correlation coefficient of the light-intensity scintillation of

beams of different wavelengths [89]. Based on the above studies, to correct the color difference in dual-wavelength free-space optical communication systems, Shubenkova experimentally confirmed the correlation between the wavefront distortion of 530-nm and 1060-nm beams [90]. Fang proposed an adaptive-optics method to correct the arrival-angle fluctuation [91].

Kibblewhite designed a deforming mirror in an adaptive-optical system according to the fluctuating power spectrum of the beam arrival angle [92]. Lukin calculated the correction accuracy and efficiency of an adaptive-optical system when the wavelength of the beacon light was inconsistent with that of the signal light [93]. Jolissaint used a Shack–Hartmann wavefront sensor to measure the wavefront-distortion results after two-color laser transmissions in the atmospheric path, and revealed the correlation between different wavelength distortions [94]. Chen analyzed the correction residual of the wavefront fluctuation of different light waves in an adaptive-optical system, and gave the correction coefficient of the correction residual value between the wavefront fluctuations of the beacon and target [95].

The key to signal wavefront-distortion correction in dual-wavelength free-space optical communication is to determine the wavefront phase relationship between the signal light and beacon light. To correct the color difference in a dual-wavelength free-space optical communication system in a weak turbulence region, a certain phase relationship was obtained between the wavefront of a dual-wavelength beam through theoretical derivation. The expression of the relationship between the distortion phase of a Gaussian beam with different wavelengths was given. Finally, experiments were conducted to verify the conclusions of the study.

Correlation Function of Arrival-Angle Fluctuation at Different Wavelengths

After Gaussian beams with respective wavelengths of λ_1 and λ_2 are transmitted through the same atmospheric channel, the variances of the arrival angles between the beams at the receiving end are respectively α_{λ_1} and α_{λ_2}. The correlation between the two can be described by a cross-correlation function [96]:

$$\langle \alpha_{\lambda_1} \alpha_{\lambda_2} \rangle = \frac{F^2}{4k_1 k_2 w_0^2} \int d^4\rho \exp\left(-\frac{2\rho^2}{w_0^2}\right) \times \left\langle \frac{\partial s_1(F, \rho_1, \lambda_1)}{\partial \rho_1} \cdot \frac{\partial s_2(F, \rho_2, \lambda_2)}{\partial \rho_2} \right\rangle T\left(\vec{\rho}\right) \tag{21.1}$$

$$T\left(\vec{\rho}\right) = \begin{cases} 1 & \rho \le D \\ 0 & \rho > D \end{cases}, \tag{21.2}$$

where $F = \frac{z}{1+(z/f)^2}$ is the curvature radius of the beam wavefront, z is the transmission distance of the beam, f is the beam parameter, $k_1 = \frac{2\pi}{\lambda_1}$ and $k_2 = \frac{2\pi}{\lambda_2}$ are the wave vectors of the beam at different wavelengths, and ω_0 is the waist radius of the beam. D is the entrance-pupil diameter of the receiving optical system. $\rho = \rho_1 - \rho_2$

represents the distance between two points on the wavefront, ρ_1 and ρ_2; when $\rho = 0$, the two points coincide.

Equation (21.1) represents the ensemble average, and its specific expression is [97]

$$\left\langle \frac{\partial s_1(F,\rho_1,\lambda_1)}{\partial \rho_1} \cdot \frac{\partial s_2(F,\rho_2,\lambda_2)}{\partial \rho_2} \right\rangle = D_S''(\rho,\lambda_1,\lambda_2)\frac{\eta^2}{\rho^2} + D_S'(\rho,\lambda_1,\lambda_2)\left(1-\frac{\eta^2}{\rho^2}\right) \tag{21.3}$$

$$D_S = (L,\rho,\lambda_1,\lambda_2) = \frac{1}{1-h^2}\left[D_S(L,\rho,\lambda) - \frac{1}{h}D_x(Lh,\rho,\lambda)\right], \tag{21.4}$$

where $D_s(\rho,\ \lambda_1,\ \lambda_2)$ is a dual-frequency phase-structure function. $h = \frac{k_1-k_2}{k_1+k_2}$, $\lambda = \frac{2\pi}{k}$, $k = \frac{2k_1k_2}{k_1+k_2}$, D_x is the logarithmic-amplitude structure function, and $D_s(L,\ \rho,\ \lambda)$ is the phase-structure function.

Equations (21.2, 21.3, and 21.4) are substituted into Eq. (21.1), and the correlation function of the arrival-angle fluctuation of Gaussian beams with different wavelengths can be written as:

$$G_\alpha(Z,\lambda_1,\lambda_2) = \sigma_\alpha^2(\lambda_1,\lambda_2) = \frac{0.101\pi^2\Gamma(1/6)}{7.963\sqrt{2}}c_n^2F^3(zR)^{-1/3} \times \frac{16}{17} \times \mathrm{Re}\left[\frac{F\left(1,1/6,23/6,\frac{1}{1+ig_1}\right)}{(1+ig_1)^{1/6}} + \frac{F\left(1,1/6,23/6,\frac{1}{1+ig_2}\right)}{(1+ig_2)^{1/6}}\right], \tag{21.5}$$

where $g_1 = \frac{F(k_1+k_2)}{k_1k_2\alpha_0^2}$, $g_2 = \frac{F(k_1-k_2)}{k_1k_2\alpha_0^2}$, $F(a,\ b,\ c,\ z)$ is the Gaussian hypergeometric function, R is the equivalent curvature radius of the wavefront, and C_n^2 is the atmospheric refractive-index structure constant.

Different Wavelength Arrival-Angle Compensation Coefficients

In a dual-wavelength free-space optical communication system, the ratio function B of the variance of the arrival-angle fluctuation of different wavelengths is defined as the arrival-angle correction coefficient of the signal light to the beacon light. When a Gaussian beam is transmitted on a near-ground horizontal path with the Kolmogorov spectrum and internal turbulence scale satisfying $l_0 \le \sqrt{\lambda L}$ in a weak turbulent region, the variance of the arrival-angle fluctuation can be expressed as [98]

$$\sigma_\alpha^2 = \frac{0.101\pi^2\Gamma(1/6)}{8\sqrt{2}}C_n^2F^3(2\alpha_0)^{-1/3}\left\{1 + \frac{16}{17}\mathrm{Re}\left[\frac{F\left(1,1/6,\frac{23}{6},\frac{1}{1+ig}\right)}{(1+ig)^{1/6}}\right]\right\}, \tag{21.6}$$

where $g = \frac{2F}{k\alpha_0^2}$. The variance-compensation coefficient B of the arrival-angle fluctuation of different wavelengths is defined as

$$B = \frac{\sigma^2{}_{\lambda_1}}{\sigma^2{}_{\lambda_2}} = \frac{\frac{0.101\pi^2\Gamma(1/6)}{8\sqrt{2}} C_n^2 F_{\lambda_1}^3 \left(2\alpha_{0_{\lambda_1}}\right)^{-1/3} \left\{1 + \frac{16}{17}\,\mathrm{Re}\left[\frac{F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_1}\right)}{\left(1+\frac{1}{1+ig_{\lambda 1}}\right)^{1/6}}\right]\right\}}{\frac{0.101\pi^2\Gamma(1/6)}{8\sqrt{2}} C_n^2 F_{\lambda_2}^3 \left(2\alpha_{0_{\lambda_2}}\right)^{-1/3} \left\{1 + \frac{16}{17}\,\mathrm{Re}\left[\frac{F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_{\lambda 2}}\right)}{\left(1+\frac{1}{1+ig_{\lambda 2}}\right)^{1/6}}\right]\right\}}, \tag{21.7}$$

where

$$B_1 = \frac{C_n^2 F_{\lambda_1}^3 \left(2\alpha_{0_{\lambda_1}}\right)^{-1/3}}{C_n^2 F_{\lambda_2}^3 \left(2\alpha_{0_{\lambda_2}}\right)^{-1/3}} \tag{21.8}$$

$$B_2 = \frac{\left\{1 + \frac{16}{17}\,\mathrm{Re}\left[\frac{F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_1}\right)}{\left(1+\frac{1}{1+ig_1}\right)^{1/6}}\right]\right\}}{\left\{1 + \frac{16}{17}\,\mathrm{Re}\left[\frac{F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_{\lambda 2}}\right)}{\left(1+\frac{1}{1+ig_{\lambda 2}}\right)^{1/6}}\right]\right\}}. \tag{21.9}$$

Then, Eq. (21.7) can be simplified as

$$B = B_1 \cdot B_2. \tag{21.10}$$

Substituting the expression of F, α_0 into Eq. (21.8) yields

$$B_1 = \frac{F_{\lambda_1}^3}{F_{\lambda_2}^3} \cdot \frac{\left(2\alpha_{0_{\lambda_1}}\right)^{-1/3}}{\left(2\alpha_{0_{\lambda_2}}\right)^{-1/3}} = \frac{\left[1 + \frac{z^2\lambda_2^2}{\pi^2\omega_0^4}\right]^{7/6}}{\left[1 + \frac{z^2\lambda_1^2}{\pi^2\omega_0^4}\right]^{7/6}}. \tag{21.11}$$

After the reverse Maclaurin-series transformation of Eq. (21.11), we obtain

$$B_1 = \left\{\frac{[2(e^{qx_2} - qx_2) + 1]}{[2(e^{x_2} - x_2) + 1]}\right\}^{6/7}, \tag{21.12}$$

where $q = \frac{\lambda_1}{\lambda_2}$ and $x_2 = \frac{z\lambda_2}{\pi\omega_0^2}$. The Gauss hypergeometric function in Eq. (21.9) is expanded into a Mellin–Barnes integral expression, and the ratio M of the two Gaussian hypergeometric functions is calculated as:

$$M = \frac{\mathrm{Re}\left[F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_1}\right)\right]}{\mathrm{Re}\left[F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_2}\right)\right]} = \int \left(\frac{g_2}{g_1}\right)^s ds. \tag{21.13}$$

We set $y = \frac{g_2}{g_1}$, according to the effective condition of a Gauss hypergeometric function on the residue integral domain. Then, the calculation result of Eq. (21.13) is

$$M = \int (y)^s ds = \frac{y-1}{\log(y)}, \tag{21.14}$$

where $y = \frac{k_1 - k_2}{k_1 + k_2}$. According to Eqs. (21.9) and (21.14), the expression of B_2 can be written as:

$$B_2 = M + \frac{1-M}{1 + \frac{16}{17}\,\mathrm{Re}\left[F\left(1, \frac{1}{6}, 23/6, \frac{1}{1+ig_2}\right)\right]}. \tag{21.15}$$

In the absence of system delays and mechanical-vibration errors, Eq. (21.7) is the proportional relation between the variance of the arrival-angle fluctuation of the signal light with wavelength λ_1 and the beacon light with wavelength λ_2 in a dual-wavelength free-space optical communication system.

Phase Relationship Between Gaussian Beams of Different Wavelengths

As shown in Fig. 21.32, the incident light is incident forward along the z axis, and the entire wavefront is distributed within the XOY plane. The wavefront measuring

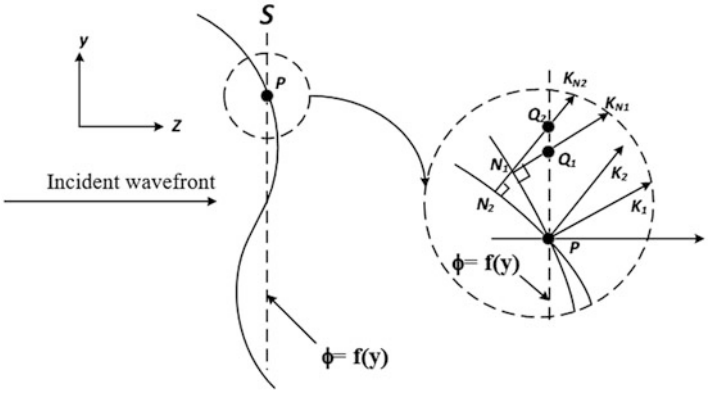


Fig. 21.32 Schematic diagram of the incident-wavefront section

surface S is perpendicular to the z axis, and the phase-distribution function of the incident light at the position of the measuring line is wave $\varphi = f(y)$.

Point P is a point on the measured wave surface. Owing to the phase distortion of the wavefront, there is a phase difference between point P on the wave surface and a nearby position. It is assumed that the wavelength of the signal light is λ_1 and the wavelength of the beacon light is λ_2. Points N_1 and N_2 are selected near point P along the equiphase plane of the two beam wavefronts, respectively. The distance between the two points PN_1 and PN_2 is ρ, and the transmission distance of the beam is L [99].

The vertical lines of the isophase plane of the beam of two wavelengths are made at points N_1 and N_2, respectively; that is, the propagation directions of the wave surface at points N_1 and N_2 are K_{λ_1} and K_{λ_2}, respectively. The two propagation vectors intersect the measuring plane S at points Q_1 and Q_2, respectively.

When ρ is small and approaches 0, triangles N_1Q_1P and N_2Q_2P can be considered as right triangles. Then, the phase difference between two points Q_1P and Q_2P is the wavefront phase difference corresponding to the beacon light and signal light, respectively. The wavefront phase difference of the corresponding signal is φ_{λ_1} and the wavefront phase difference of the beacon is φ_{λ_2}. Under the above conditions, the wavefront-distortion phases of the signal light and beacon light can be expressed as

$$\varphi_{\lambda_2} = \varphi_p - \varphi_Q = \frac{2\pi}{\lambda_2} N_2Q_2 = \frac{2\pi}{\lambda_2}\left(n_p - n_Q\right)L \tag{21.16}$$

$$\varphi_{\lambda_1} = \varphi_p - \varphi_{Q1} = \frac{2\pi}{\lambda_1} N_1Q_1 = \frac{2\pi}{\lambda_1}\left(n_p - n_{Q1}\right)L. \tag{21.17}$$

Because the signal light and beacon light are transmitted along the same path, the refractive index of air varies the same between corresponding points, so

$$n_p - n_{Q1} = n_p - n_{Q2}. \tag{21.18}$$

When the signal light and beacon light are transmitted in the same atmospheric channel, the random variation of the refractive index in atmospheric turbulence has different modulation effects on Gaussian beams of different wavelengths. There is also some difference between the wavefront fluctuation distortion; that is, the phase difference between PQ_2 and PQ_1 is different, and the corresponding wavefront-distortion phase difference is

$$\begin{aligned} \Delta &= \left|\varphi_{\lambda_1} - \varphi_{\lambda_2}\right| \\ &= \left|\left(\varphi_p - \varphi_Q\right) - \left(\varphi_p - \varphi_{Q1}\right)\right| \\ &= 2\pi\left(n_p - n_Q\right)L\left|\frac{\lambda_1 - \lambda_2}{\lambda_1\lambda_2}\right|. \end{aligned} \tag{21.19}$$

Then, the relationship between the phase of the signal-light front φ_{λ_1} and the phase of the beacon-light front φ_{λ_2} at point P is

$$\varphi_{\lambda_1} = \frac{\frac{2\pi}{\lambda_2}(n_p - n_Q)L \pm \Delta}{\frac{2\pi}{\lambda_2}(n_p - n_Q)L} \cdot \varphi_{\lambda_2}. \tag{21.20}$$

Equations (21.18) and (21.19) are substituted into Eq. (21.20), and the relationship between the phase distortion of Gaussian beams with different wavelengths is as follows:

$$\varphi_{\lambda_1} = \left(\frac{\lambda_2}{\lambda_1} \pm \left|\frac{\lambda_1 - \lambda_2}{\lambda_1 \lambda_2}\right|\right)\varphi_{\lambda_2}. \tag{21.21}$$

According to this relation, the wavefront-distortion phase difference of the beam with different wavelengths can be calculated, based on the known wavelength of the signal light and beacon light, the transmission distance L of the beam, and the change of refractive index. The wavefront-distortion measurement value of the signal light can also be calculated according to the wavefront-distortion measurement value of the beacon light.

Wavefront-Distortion Correction for Dual-Wavelength Free-Space Optical Communication Systems

The correction matrix is calculated according to the quantitative relationship between the wavefront-distortion phase of a Gaussian beam with a different wavelength, and the wavefront-reconstruction matrix of the beacon is corrected, so that the wavefront-reconstruction matrix of the signal light is obtained to correct the signal light. The distorted wavefront $W(x, y)$ can be expressed as follows:

$$W(x,y) = W^{'}(x,y) - W_0(x,y) = \sum_{j=0}^{j=\max} W_j Z_j(x,y), \tag{21.22}$$

where Z_j is a Zernike polynomial, W_j is the coefficient of the Zernike polynomial, and j is the order of the Zernike polynomial. $W(x, y)$ is used to measure the wavefront and $W_0(x, y)$ represents the reference-plane wavefront. The difference between the two produces the distorted wavefront. A Gaussian beam with wavelength λ_2 is emitted as a common optical path of the beacon light and the signal light with wavelength λ_1. A Shack–Hartmann wavefront sensor measures the beacon-light wavefront to obtain the beacon-light gradient matrix containing the distortion:

$$\begin{bmatrix} b_{\lambda_2}(x,y)|_1 \\ b_{\lambda_2}(x,y)|_2 \\ \vdots \\ b_{\lambda_2}(x,y)|_k \\ c_{\lambda_2}(x,y)|_1 \\ c_{\lambda_2}(x,y)|_2 \\ \vdots \\ c_{\lambda_2}(x,y)|_k \end{bmatrix} = \begin{bmatrix} g_1(x,y)|_1 & g_2(x,y)|_1 & \cdots & g_{j\max}(x,y)\Big|_1 \\ g_1(x,y)|_2 & g_2(x,y)|_2 & \cdots & g_{j\max}(x,y)\Big|_2 \\ & & \vdots & \\ g_1(x,y)|_k & g_2(x,y)|_k & \cdots & g_{j\max}(x,y)\Big|_k \\ h_1(x,y)|_1 & h_2(x,y)|_1 & \cdots & h_{j\max}(x,y)\Big|_1 \\ h_1(x,y)|_2 & h_2(x,y)|_2 & \cdots & h_{j\max}(x,y)\Big|_2 \\ & & \vdots & \\ h_1(x,y)|_k & h_2(x,y)|_k & \cdots & h_{j\max}(x,y)\Big|_k \end{bmatrix} \cdot \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_{j\max} \end{bmatrix}, \tag{21.23}$$

where $\frac{\Delta x(x,y)}{f}_{\lambda_2}\Big|_k = b_{\lambda_2}(x,y)|_k$ and $\frac{\Delta y(x,y)}{f}_{\lambda_2}\Big|_k = c_{\lambda_2}(x,y)|_k$. According to Eq. (21.21), we calculate the offset gradient of the Gaussian-beam distortion with two different wavelengths on the sensor measurement surface, and the following relation can be obtained:

$$\frac{\Delta x(x,y)}{f}_{\lambda_1}\Big|_k = \frac{1}{\gamma} \cdot \frac{\Delta x(x,y)}{f}_{\lambda_2}\Big|_k = b_{\lambda_1}(x,y)|_k \tag{21.24}$$

$$\frac{\Delta y(x,y)}{f}_{\lambda_1}\Big|_k = \frac{1}{\gamma} \cdot \frac{\Delta y(x,y)}{f}_{\lambda_2}\Big|_k = c_{\lambda_1}(x,y)|_k, \tag{21.25}$$

where $\frac{1}{\gamma} = \frac{\lambda_2}{\lambda_1} \cdot \left(\cos A\theta_2 + \frac{\sin A\theta_2}{\tan\theta_2}\right)$. According to this relation, a coefficient correction matrix can be obtained:

$$\gamma = \begin{bmatrix} \gamma_1 & 0 \cdots 0_k \\ 0 & \gamma_2 \cdots 0_k \\ & \ddots \\ 0\ 0 & \cdots \gamma_k \end{bmatrix}. \tag{21.26}$$

The wavefront-reconstruction matrix of the beacon light is obtained by multiplying it by the coefficient-correction matrix. In the AO control system, beacon-light wavefront-reconstruction matrix and beacon-light wavefront-reconstruction matrix obtained after correction are, respectively, used to correct the signal-light wavefront distortion. The signal-light Strehl ratio (SR) and three-dimensional light spot of the beam are obtained, as shown in Fig. 21.33.

Figure 21.33 shows the SR-value change curve of the system in the correction process. (a) is the SR curve when the distorted beacon-light wavefront is directly used to correct the signal-light wavefront distortion. (b) is the SR curve when the

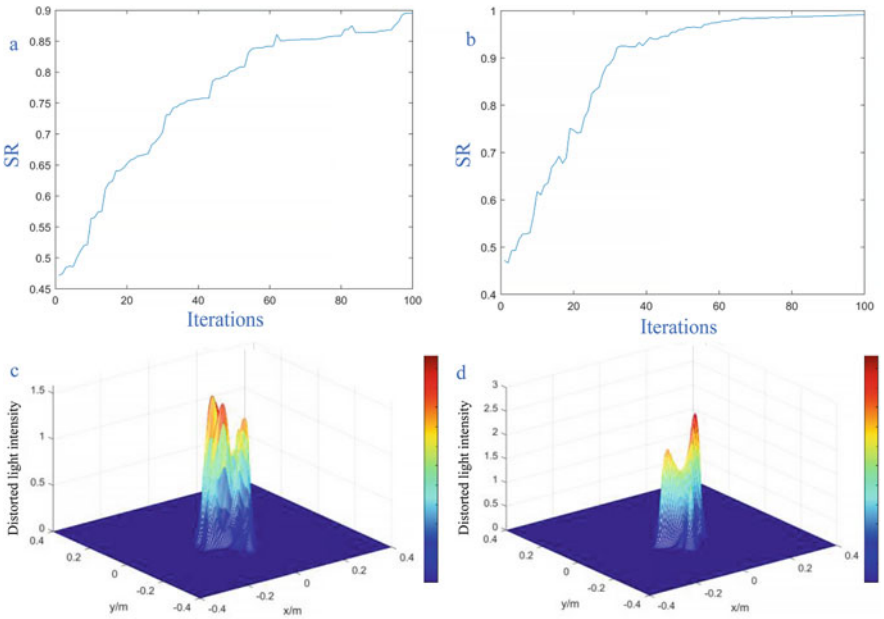


Fig. 21.33 Distorted beacon-light front. Corrected distorted signal-light front. (**a**, **c**) Without the correction coefficient; (**b**, **d**) with the wavefront-correction coefficient

wavefront distortion of the beacon is corrected by the wavefront phase matrix modified by Eq. (21.21) to correct the wavefront distortion of the signal.

As can be seen from the figure, the corrected SR ratio of the wavefront reconstructed matrix after correction converges to 1, and the overall SR ratio increases by 10%; the wavefront reconstructed matrix converges to 0.9 before correction. When correcting the signal light after correcting the beacon-light wavefront-distortion matrix, the SR of the system converges quickly, and tends to stabilize when the number of iterations is 80. The jitter of the SR is obviously reduced, which verifies the correctness of the phase relationship between the signal light and the beacon-light wavefront derived from Eq. (21.21).

Experimental Results and Analysis

The experimental schematic diagram is shown in Fig. 21.34. The signal light and beacon light are transmitted through the common optical path of the beam synthesizer, then through the same atmospheric channel, reflected by the plane mirror, and transmitted to the receiving antenna on the fifth and seventh floors.

After the beam combination, the beam is reflected by the deformable mirror and the reflector and then incident on the beam-splitter mirror. The beam-splitter mirror separates the signal light and beacon light, and the wavefront sensor, respectively, detects the wavefront distortion of the beacon light and signal light. After adaptive-

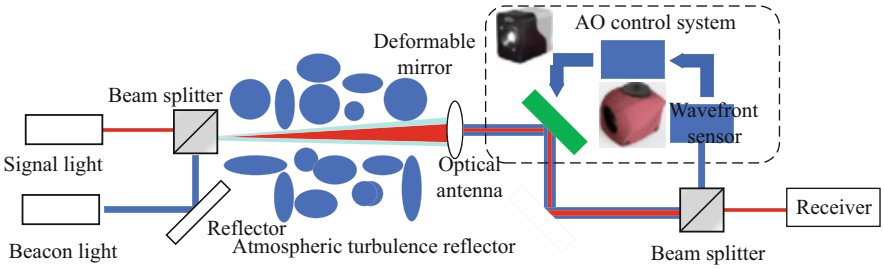


Fig. 21.34 Schematic diagram of experimental structure

Table 21.3 Parameters of the experimental equipment

Equipment	Model number	Parameter
Wavefront sensor	THORLABS-HASO4-NIR	Wavelength range: 1500–1600 nm Subaperture parameter: 32 × 40
Wavefront sensor	THORLABS-HASO4-FIRST	Wavelength range: 400–1100 nm Subaperture number: 32 × 40
Beacon laser light source	MW-GX-650	Wavelength: 650 nm, linewidth: 0–3 nm, output power: 35 mW
Signal laser source	KOHERA-BASIK Module	Wavelength: 1550 nm, linewidth: 0.1 kHz, output power: 0–200 mW
Deformable mirror	ALPAO-DM	Number of deformed mirrors: 69, effective aperture: 10.5 mm Frequency: 800 Hz
Optical antenna	Maksutov–Cassegrain	Diameter: 25 mm, focal length: 300 mm
Laser collimator	F810APC	Wavelength: 1550 nm/650 nm, focal length: 37.13 mm

PID algorithm control, the optical wavefront-distortion of the signal is corrected [19]. Table 21.3 shows the corresponding experimental instrument parameters.

After the AO system captures the beacon light, the wavefront detector measures its gradient matrix, generates the wavefront-reconstruction matrix, and uses it to correct the signal optical-wavefront distortion. The wavefront peak–valley (PV) value and the wavefront phase-fluctuation RMS value of the signal light before and after correction are shown in Figs. 21.35 and 21.36, respectively.

The mean value of the wavefront PV value of the signal light is 2.08 μm and the variance of the PV value is 0.90 μm^2 after directly correcting the signal-light distortion data. The mean RMS of the wavefront fluctuation is 0.40 μm^2, and the variance of the wavefront fluctuation is 0.04 μm^2. Then, the wavefront-reconstruction matrix of the beacon light is multiplied by the coefficient-correction matrix to obtain the wavefront-reconstruction matrix of the signal light. The AO system corrects the wavefront of the signal light according to its wavefront-reconstruction matrix, and obtains the wavefront PV value and wavefront phase-fluctuation RMS value of the signal light, as shown in Fig. 21.36.

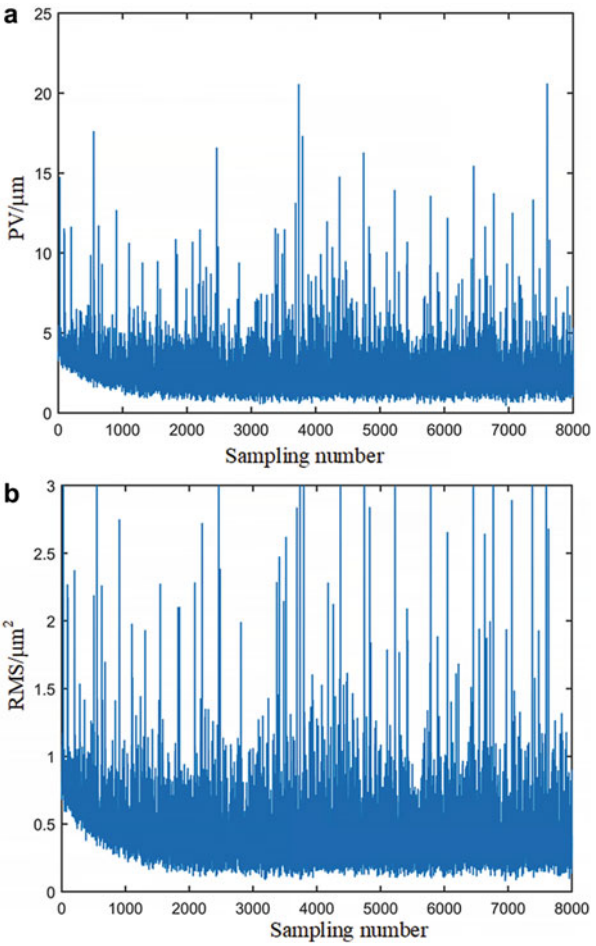


Fig. 21.35 Beacon-light correction signal. (**a**) Wavefront PV-value change curve; (**b**) change curve of wavefront RMS value

The mean PV value and variance of the signal wavefront are 2.05 μm and 0.46 μm^2, respectively, after correcting the signal wavefront using the beacon light–corrected wavefront-reconstruction matrix. The mean RMS of the wavefront phase fluctuation is 0.35 μm^2, and the variance of the RMS of the wavefront phase fluctuation is 0.01 μm^2. Compared with the wavefront-reconstruction matrix, the variance of the wavefront PV value decreased by 49.05% and the variance of the wavefront phase-fluctuation RMS decreased by 35.58% in the signal optical-reconstruction matrix correction mode.

The color difference of an adaptive optical system is a systematic error caused by the inconsistency between the wavelength of the system signal light and that of the beacon light. In this chapter, the correlation between the arrival-angle fluctuations of different wavelengths is analyzed. Moreover, the correlation between the number function of the correlation between the arrival-angle fluctuations of Gaussian beams

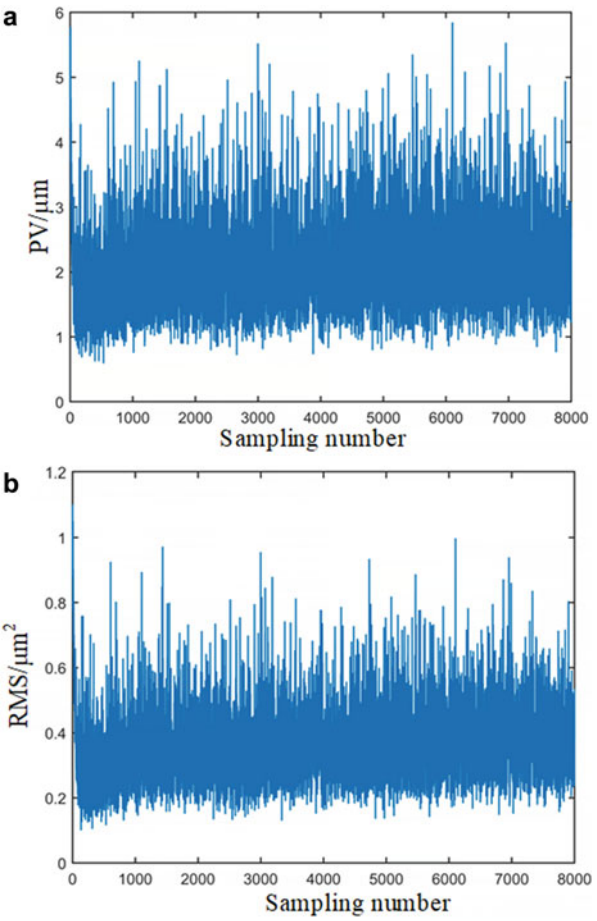


Fig. 21.36 Beacon light–corrected wavefront-reconstruction matrix correcting the signal light. (**a**) Change curve of the wavefront PV value; (**b**) change curve of the wavefront RMS value

of different wavelengths and the wavefront phase of distorted Gaussian beams of different wavelengths is deduced. The effects of the transmission distance, atmospheric refraction-index structure constant, Gaussian beam-waist radius, and wavelength difference on the correlation number of the arrival-angle fluctuation are analyzed by simulation.

The results show that the mutual correlation between the arrival-angle fluctuations of different wavelengths increases with the increase of the transmission distance. With the increase of the refraction-index structure constant, the correlation gradually decreases. The correlation is mainly affected by the wavelength and is negatively correlated with the wavelength difference between the signal light and beacon light. The smaller the beam waist radius is, the more the correlation between the arrival-angle fluctuation of different wavelengths decreases, the smaller the

wavelength difference is, and the smaller the difference between the distortion phase of the Gaussian beam of a different wavelength is.

The correction-coefficient matrix of the wavefront-reconstruction matrix is obtained according to the relationship between the phases of different-wavelength beams, and it is applied to the simulation correction experiment. The simulation experiment results show that the SR value of the signal light is improved by 10% by using the modified wavefront-reconstruction matrix for signal-light correction.

In the field experiment, the PV variance and RMS variance of the signal light were 0.9 μm^2 and 0.046 μm^2, respectively, without using the wavefront-reconstruction matrix correction coefficient. The wavefront PV variance and RMS variance of the signal light were 0.46 μm^2 and 0.012 μm^2, respectively, when correcting using the modified wavefront-reconstruction matrix. The PV variance and RMS variance of the signal light were reduced by 49.05% and 35.58%, respectively.

Coherent-Detection Performance Optimization

Influence of the Polarization State of Partially Coherent Light on the Sensitivity of Coherent Detection in Non-Kolmogorov Atmospheric Turbulence

Theoretical Analysis

Compared with traditional incoherent optical communication, spatial coherent optical communication can increase its receiving sensitivity by 20–23 dB and allow longer transmission distances under the same transmission power, which is more conducive to the suppression of atmospheric turbulence and channel fading [1]. When a beam is transmitted in an atmospheric channel, the atmospheric flicker, arrival-angle fluctuation, and beam drift will occur, owing to the influence of the random fluctuation of the atmospheric refractive index, which significantly reduces the performance of the optical wireless communication link. It was found that, compared with fully coherent light, partially coherent light was less affected by turbulence and had stronger stability [100].

The sensitivity of a coherent optical communication system is often described using the performance parameters of the detector; however, the performance of the detector is not the only determining factor, it is also affected by other parts of the system, such as granular noise, phase noise, local light power, mixing efficiency, etc. The influence of the former is main and direct; the influence of the latter is secondary and indirect. These factors have been the subjects of intense research.

For a sensitivity analysis of a heterodyne-detection system of space completely coherent light, Abbas and Slusher et al. [101, 102] studied the influence of scattered particle noise caused by local vibration light on the sensitivity of the coherent-detection system. He et al. [103] studied the noise-suppression mechanism of coherent optical balance detection, which effectively suppressed the influence of the noise fluctuation on the system output and improved the sensitivity.

Nan hang et al. [104] studied the influence of the loss of support between the local light and signal light on the detection sensitivity of spatially coherent optical communication under atmospheric turbulence. Liang et al. [105] studied the influence of local vibration power on the coherent-detection sensitivity of a space balanced detector.

Most physically realizable light sources radiate randomly, owing to inhomogeneity in the laser resonator and the spontaneous emission of atoms. In other words, the actual radiation is a partially coherent light beam; hence, it is of great significance to study the sensitivity of a heterodyne system of partially coherent light beams.

In recent years, the performance of partially coherent light for heterodyne detection has been extensively studied. Tanaka et al. [106] calculated the heterodyne efficiency for heterodyne detection when both the signal light and local oscillator light were partially coherent beams. Pu et al. [107] studied the system performance of partially coherent laser beams used in the heterodyne detection of lidar. Salem and Rolland et al. [108] studied the effect of the angular error between the signal light and local oscillator light of a partially coherent beam on the performance of a coherent-detection system. Li et al. [11] discussed the influence of turbulence on a heterodyne-detection system, based on the theory of partially coherent beams.

The above work is based on the scalar theory of partially coherent light fields. In addition, the polarization phenomenon of the electromagnetic field cannot be ignored, and the polarization of the beam usually changes with the propagation of the beam. Salem et al. [109] studied the influence of coherence and polarization changes on the heterodyne detection of random electromagnetic beams propagating in free space. Wang et al. [110] studied the coherent polarization matrix of a Gaussian–Schell electromagnetic beam under a horizontal link and deduced the analytical expression of the heterodyne-detection system.

In the above research work, the influence of atmospheric turbulence on the heterodyne efficiency of a coherent optical communication system is based on Kolmogorov turbulence. In fact, atmospheric turbulence greatly deviates from the turbulence described by the Kolmogorov power spectrum, while non-Kolmogorov turbulence is closer to the actual situation of atmospheric turbulence.

At present, little research has been conducted on the influence of the beam-polarization state on the sensitivity of coherent detection when a partially coherent beam is transmitted up and down in non-Kolmogorov atmospheric turbulence. Based on the above research, we derived the expression of the sensitivity of a heterodyne-detection system when partially correlated light was transmitted in an oblique path under the non-Kolmogorov turbulence spectrum. The effects of the power-spectrum power law, turbulence scale, transmission distance, receiving height, and detector diameter on the sensitivity of an electromagnetic Gaussian–Schell model (EGSM) beam transmitted through uplinks and downlinks under non-Kolmogorov atmospheric turbulence were analyzed, and meaningful results were obtained.

A partially coherent Gauss–Schell model (GSM) is a scalar light field that can be considered to be completely polarized, regardless of its vector properties [111]. However, the light field is usually not absolutely polarized, but partial polarization

of the electromagnetic field occurs, in the form of partially polarized light. A partially coherent EGSM beam is a vector GSM beam, and its polarization is considered.

In this study, an EGSM beam is taken as the research object, and the system performance of an EGSM beam for heterodyne detection under non-Kolmogorov turbulence is discussed. As shown in Fig. 21.37, it is considered that both the signal light and local oscillator light are partially coherent EGSM beams, and frequency mixing occurs on the detector photosensitive surface at $z = 0$. k_o and k_s are wave vectors of the local and signal light, respectively. Assuming that both the signal light and local oscillator light are parallel to the optical axis and perpendicular to the detector surface, the two light fields on the detector surface can be expressed as [112]

$$U_o(\boldsymbol{\rho}, t) = \left[U_{ox}(\boldsymbol{\rho})\boldsymbol{e}_x + U_{oy}(\boldsymbol{\rho})\boldsymbol{e}_y\right] \exp(j\omega_o t) \tag{21.27}$$

$$U_s(\boldsymbol{\rho}, t) = \left[U_{sx}(\boldsymbol{\rho})\boldsymbol{e}_x + U_{sy}(\boldsymbol{\rho})\boldsymbol{e}_y\right] \exp(j\omega_s t), \tag{21.28}$$

where $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$ are unit vectors, and ω_o and ω_s are the angular frequencies of the local light and signal light, respectively.

The polarization properties and spatial-coherence properties of partially coherent beams can be represented by beam-coherence polarization (BCP) matrix J_s. The two types of signal transmission in the z direction of the vibration signal light and the light beam of BCP matrix $J_s(\rho_1, \rho_2)$ and $J_o(\rho_1, \rho_2)$ can be, respectively, represented as [112]

$$J_s(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) = \begin{bmatrix} \left\langle U_{sx}^*(\boldsymbol{\rho}_1) U_{sx}(\boldsymbol{\rho}_2) \right\rangle & \left\langle U_{sx}^*(\boldsymbol{\rho}_1) U_{sy}(\boldsymbol{\rho}_2) \right\rangle \\ \left\langle U_{sy}^*(\boldsymbol{\rho}_1) U_{sx}(\boldsymbol{\rho}_2) \right\rangle & \left\langle U_{sy}^*(\boldsymbol{\rho}_1) U_{sy}(\boldsymbol{\rho}_2) \right\rangle \end{bmatrix} \tag{21.29}$$

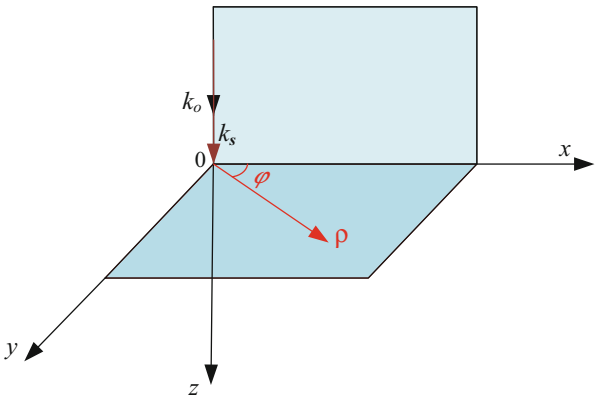


Fig. 21.37 Mixing model of a beam on the detector in the $z = 0$ plane

$$J_o(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2) = \begin{bmatrix} \left\langle U_{ox}^*(\boldsymbol{\rho}_1)U_{ox}(\boldsymbol{\rho}_2)\right\rangle & \left\langle U_{ox}^*(\boldsymbol{\rho}_1)U_{oy}(\boldsymbol{\rho}_2)\right\rangle \\ \left\langle U_{oy}^*(\boldsymbol{\rho}_1)U_{ox}(\boldsymbol{\rho}_2)\right\rangle & \left\langle U_{oy}^*(\boldsymbol{\rho}_1)U_{oy}(\boldsymbol{\rho}_2)\right\rangle \end{bmatrix}, \tag{21.30}$$

where ρ_1 and ρ_2 represent the position vector of any two points on the receiving end face. According to the heterodyne-detection principle and the double-balance detection principle, the power of the IF signal is obtained as [112]

$$P_{IF} = 2\Re^2 \iint\iint \mathrm{Re}\{\mathrm{Tr}[J_o(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)J_s(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)]\}d^2\boldsymbol{\rho}_1 d^2\boldsymbol{\rho}_2, \tag{21.31}$$

where Tr is the type of matrix trace, $\Re = e\eta/h\nu$ is the response of the detector, e is the electron charge, η is the quantum efficiency, and $h\nu$ is the photon energy. Here, it is convenient for analysis that the response of the detector $\Re$ is evenly distributed on the surface of the photosensitive detector. The main noise source in the coherent-detection system is the loose particle noise generated by the local vibration light, which can be expressed as [112]

$$P_n = eB\Re \iint \mathrm{Tr}[J_o(\boldsymbol{\rho},\boldsymbol{\rho})]d^2\boldsymbol{\rho}, \tag{21.32}$$

where B is the bandwidth of the detector. By definition, the signal-to-noise ratio of the coherent-detection system for partially coherent EGSM beams is [112]

$$SNR_{EGSM} = \frac{2\Re \iint\iint \mathrm{Re}\{\mathrm{Tr}[J_o(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)J_s(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)]\}d^2\boldsymbol{\rho}_1 d^2\boldsymbol{\rho}_2}{eB\iint \mathrm{Tr}[J_o(\boldsymbol{\rho},\boldsymbol{\rho})]d^2\boldsymbol{\rho}}. \tag{21.33}$$

According to the definition of heterodyne efficiency [112], the expression of the heterodyne efficiency of a partially coherent EGSM beam for heterodyne detection can be given as:

$$\eta_{EGSM} = \frac{\iint\iint \mathrm{Re}\{\mathrm{Tr}[J_o(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)J_s(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2)]\}d^2\boldsymbol{\rho}_1 d^2\boldsymbol{\rho}_2}{\iint \mathrm{Tr}\left[J_o(\boldsymbol{\rho},\boldsymbol{\rho})d^2\boldsymbol{\rho}\right]\iint \mathrm{Tr}\left[J_s(\boldsymbol{\rho},\boldsymbol{\rho})d^2\boldsymbol{\rho}\right]}. \tag{21.34}$$

The performance of the system when an EGSM beam is used for heterodyne detection under turbulent conditions will now be analyzed based on the EGSM beam. Considering an electromagnetic Gaussian–Schell model beam radiated by a Gaussian–Schell source, its BCP matrix is [113]

$$J_{\alpha ij}^{(0)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = I_{\alpha ij}\exp\left[-\frac{{\boldsymbol{r}_1}^2+{\boldsymbol{r}_2}^2}{4w_\alpha^2} - \frac{(\boldsymbol{r}_1-\boldsymbol{r}_2)^2}{2\delta_{\alpha ij}^2}\right], \tag{21.35}$$

where r_1 and r_2 represent the position vectors of any two points at the source field $z = 0$; $I_{\alpha ij}$ as the light intensity, w_α is the radius of the beam, $\delta_{\alpha ij}$ is the coherence

length of the space, and the subscript α for s or o, respectively, corresponds to the vibration signal light and the light, $i, j = x, y$.

In atmospheric turbulence, the EGSM beam is transmitted from the source plane $z = 0$ to $z > 0$. According to the generalized Huygens–Fresnel principle, the expression of each element of the BCP matrix of the beam at the receiving end is [114]

$$\begin{aligned} J_{\alpha ij}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, z_\alpha) = I_{\alpha ij}\left(\frac{k}{2\pi z_\alpha}\right)^2 \iint d^2\boldsymbol{r}_1 \iint d^2\boldsymbol{r}_2 \times \exp\left[-ik\frac{(\boldsymbol{\rho}_1 - \boldsymbol{r}_1)^2 - (\boldsymbol{\rho}_2 - \boldsymbol{r}_2)^2}{2z_\alpha}\right] \\ \times \exp\left[-\frac{\boldsymbol{r}_1{}^2 + \boldsymbol{r}_2{}^2}{4w_\alpha^2} - \frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2)^2}{2\delta_{\alpha ij}^2}\right] \\ \times < \exp[\psi^*(\boldsymbol{\rho}_1, \boldsymbol{r}_1, z_\alpha) + \psi(\boldsymbol{\rho}_2, \boldsymbol{r}_2, z_\alpha)] >_m, \end{aligned} \tag{21.36}$$

where z_α is the beam-propagation distance; k = 2π/λ, the spatial wave number for freedom for wavelength λ; * is the complex conjugate;$< >_m$ is the ensemble average of the turbulent medium, and $\psi(\boldsymbol{\rho}, \boldsymbol{r}, z_\alpha)$ is a complex turbulent-medium random disturbance.

$< \exp [\psi^*(\boldsymbol{\rho}_1, \boldsymbol{r}_1, z_\alpha) + \psi(\boldsymbol{\rho}_2, \boldsymbol{r}_2, z_\alpha)]>_m$, which describes the effect of turbulence, can be expressed as [115]

$$\begin{aligned} &< \exp[\psi^*(\boldsymbol{\rho}_1, \boldsymbol{r}_1, z_\alpha) + \psi(\boldsymbol{\rho}_2, \boldsymbol{r}_2, z_\alpha)] >_m \\ &= \exp\left[-\frac{1}{2}D_\phi(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2, \boldsymbol{r}_1 - \boldsymbol{r}_2)\right]. \end{aligned} \tag{21.37}$$

In this formula, $D_\phi(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2, \boldsymbol{r}_1 - \boldsymbol{r}_2)$ is the wave-structure function that transmits the signal instantaneously to the receiving end [116]:

$$\begin{aligned} &< \exp[\psi^*(\boldsymbol{\rho}_1, \boldsymbol{r}_1, z_\alpha) + \psi(\boldsymbol{\rho}_2, \boldsymbol{r}_2, z_\alpha)] >_m \\ &\cong \exp\left[-\frac{1}{2}\left(M_1\boldsymbol{r}_d^2 + M_2\boldsymbol{r}_d\boldsymbol{\rho}_d + M_3\boldsymbol{\rho}_d^2\right)\right], \end{aligned} \tag{21.38}$$

where $\boldsymbol{\rho}_d = \boldsymbol{\rho}_1 - \boldsymbol{\rho}_2$; $\boldsymbol{r}_d = \boldsymbol{r}_1 - \boldsymbol{r}_2$; M_1, M_2, and M_3 are

$$\begin{cases} M_1 = 2\pi^2 k^2 \sec\varphi \int_{h_0}^{H}\int_0^\infty \Phi_n(\kappa, h)(1-\xi)^2\kappa^3 d\kappa dh \\ M_2 = 4\pi^2 k^2 \sec\varphi \int_{h_0}^{H}\int_0^\infty \Phi_n(\kappa, h)(1-\xi)\xi\kappa^3 d\kappa dh \\ M_3 = 2\pi^2 k^2 \sec\varphi \int_{h_0}^{H}\int_0^\infty \Phi_n(\kappa, h)\xi^2\kappa^3 d\kappa dh \end{cases}, \tag{21.39}$$

where, in the uplink, $\xi = 1 - (h - h_0) / (H - h_0)$, H is the height of the receiving end, and h_0 is the height of the light source. In the downlink, $\xi = (h - h_0) / (H - h_0)$, H is the height of the light source, and h_0 is the height of the receiving end. φ is the zenith angle. For $\Phi_n(\kappa, h)$, a non-Kolmogorov turbulence spectrum is adopted, and its expression is [117]

$$\Phi_n(\kappa, h, \alpha) = A(\alpha)\widetilde{C}_n^2 \frac{\exp\left[-\kappa^2/\kappa_m^2\right]}{\left(\kappa^2 + \kappa_0^2\right)^{\alpha/2}}, 0 \le \kappa < \infty, 3 < \alpha < 4. \tag{21.40}$$

For non-Kolmogorov turbulence, $\int_0^\infty \Phi_n(\kappa, h, \alpha)\kappa^3 d\kappa$; after integration, we can obtain

$$\int_0^\infty \Phi_n(\kappa, h, \alpha)\kappa^3 d\kappa = \frac{A(\alpha)}{2(\alpha - 2)}\widetilde{C}_n^2 \left[\kappa_m^{2-\alpha}\beta \exp\left(\frac{\kappa_0^2}{\kappa_m^2}\right)\Gamma\left(2 - \frac{\alpha}{2}, \frac{\kappa_0^2}{\kappa_m^2}\right) - 2\kappa_0^{4-\alpha}\right], \tag{21.41}$$

where $\beta = 2\kappa_o^2 - 2\kappa_m^2 + \alpha\kappa_m^2$, α is the power-spectrum power law; $\kappa_0 = 2\pi/L_0$, L_0 is the outer turbulence scale; $A(\alpha) = \Gamma(\alpha - 1)\cos(\alpha\pi/2)/(4\pi^2)$, $\Gamma(x)$ indicates the gamma function; and $\kappa_m = c(\alpha)/l_0$, l_0 is a turbulent scale in which $c(\alpha)$ is a function of the power-spectrum power law α, with the following type [119]:

$$c(\alpha) = \left[\Gamma\left(5 - \frac{\alpha}{2}\right)A(\alpha)\frac{2}{3}\pi\right]^{1/(\alpha-5)}, \tag{21.42}$$

where $\widetilde{C}_n^2$ is the Kolmogorov-turbulence atmospheric refractive-index structure constant, with unit $m^{3-\alpha}$, proposed by the International Telecommunication Union—Radiocommunication Sector (ITU-R), and changes with the height of the atmospheric refractive-index structure constant model; its expression is as follows [118]:

$$\begin{aligned} C_n^2(h) = {} & 8.148 \times 10^{-56} v_{RMS}^2 h^{10} e^{-h/1000} + 2.7 \times 10^{-16} e^{-h/1500} \\ & + C_0 e^{-h/100}, \end{aligned} \tag{21.43}$$

where

$$v_{RMS} = \sqrt{v_g^2 + 30.69 v_g + 348.91}, \quad \widetilde{C}_n^2 = \gamma C_n^2,$$

where γ is a constant 1, v_{RMS} is the vertical-path wind speed, v_g is the near-surface wind speed, and v_g is approximately 2.8 m/s when the surface wind speed is unknown. C_0 represents the atmospheric structure constant near the surface.

By substituting Eq. (21.37) into Eq. (21.36) and integrating, we can obtain

$$J_{\alpha ij}(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,z_\alpha)=\frac{I_{\alpha ij}}{Q_{\alpha ij}}\exp\left[-H_{\alpha ij}(\boldsymbol{\rho}_1-\boldsymbol{\rho}_2)^2\right]$$
$$\times\ \exp\left[-\frac{1}{8w_\alpha^2 Q_{\alpha ij}}(\boldsymbol{\rho}_1+\boldsymbol{\rho}_2)^2+ikT_{\alpha ij}\left({\boldsymbol{\rho}_2}^2-{\boldsymbol{\rho}_1}^2\right)\right],\qquad(21.44)$$

where

$$Q_{\alpha ij}=1+\left(\frac{z_\alpha}{kw_\alpha\theta_{\alpha ij}}\right)^2+\frac{{z_\alpha}^2M_3}{k^2w_\alpha^2}$$

$$H_{\alpha ij}=\frac{1}{2\theta_{\alpha ij}^2Q_{\alpha ij}}+\frac{1}{Q_{\alpha ij}}\left(\frac{M_3}{2}+\frac{M_2}{2}\right)+\frac{M_1}{2}-\frac{\left(\frac{M_2}{2}\right)^2{z_\alpha}^2}{2w_\alpha^2k^2Q_{\alpha ij}}\qquad(21.45)$$

$$\frac{1}{\theta_{\alpha ij}^2}=\frac{1}{4w_\alpha^2}+\frac{1}{\delta_{\alpha ij}^2}$$

$$T_{\alpha ij}=\frac{k^2w_\alpha^2Q_{\alpha ij}+\left(\frac{M_2}{2}\right){z_\alpha}^2-k^2w_\alpha^2}{2k^2z_\alpha w_\alpha^2Q_{\alpha ij}}.$$

For a random electromagnetic beam, the polarization degree is defined as [115]

$$P(\boldsymbol{\rho},\boldsymbol{\rho},\mathrm{z})=\sqrt{1-\frac{4\mathrm{Det}J(\boldsymbol{\rho},\boldsymbol{\rho},\mathrm{z})}{[TrJ(\boldsymbol{\rho},\boldsymbol{\rho},\mathrm{z})]^2}},\qquad(21.46)$$

where *Det* and *Tr*, respectively, represent the determinant and trace of matrix *J*.

By substituting Eq. (21.44) into Eq. (21.46), the polarization degree of the point on the beam axis in the receiving plane at $z>0$ is obtained as:

$$P(\boldsymbol{\rho}=0,\mathrm{z})=\sqrt{1-\frac{4\left(\frac{I_{\alpha xx}I_{\alpha yy}}{Q_{\alpha xx}Q_{\alpha yy}}-\frac{I_{\alpha xy}I_{\alpha yx}}{Q_{\alpha xy}Q_{\alpha yx}}\right)}{\left(\frac{I_{\alpha xx}}{Q_{\alpha xx}}+\frac{I_{\alpha yy}}{Q_{\alpha yy}}\right)^2}}.\qquad(21.47)$$

The degree of polarization is $0\le P\le 1$. When $P=1$, the beam is fully polarized. When $0<P<1$, the beam is partially polarized. When $P=0$, the beam is unpolarized.

We substitute Eq. (21.44) into Eq. (21.31), and convert the operation into a polar coordinate system. The *IF* signal output by the detector can be expressed as

$$P_{IFij}=2\Re^2\int_0^{2\pi}\int_0^{2\pi}\int_0^{D/2}\int_0^{D/2}\mathrm{Re}\left[J_{oij}(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,z_\alpha)\times {J_{sij}}^*(\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,z_\alpha)\right]\rho_1\rho_2d\rho_1d\rho_2d\varphi_1d\varphi_2,\qquad(21.48)$$

where D is the diameter of the effective detector area, the Cartesian coordinate system is established with the center of the detector as the origin, and φ is the included angle of the position vector $\boldsymbol{\rho}$ and the x-axis.

To facilitate an integral calculation, a Gaussian-type aperture or soft aperture is considered to replace a hard aperture here. The relation between the soft aperture W and hard aperture D is [115] $D^2 = 8\ W^2$. Based on the relation between the soft aperture and hard aperture, Eq. (21.48) can be arranged as follows:

$$\begin{aligned} P_{IFij} = 2\Re^2 \frac{I_{oij} I_{sij}}{Q_{oij} Q_{sij}} \int_0^{2\pi} \int_0^{2\pi} \int_0^{\infty} \int_0^{\infty} \mathrm{Re}\left\{ \frac{I_{oij}}{Q_{oij}} \exp\left[-H_{oij}(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)^2\right] \exp\left[-\frac{1}{8w_o^2 Q_{oij}}(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2)^2\right] \right. \\ \times \exp\left[-ikT_{oij}\left(\boldsymbol{\rho}_1^2 - \boldsymbol{\rho}_2^2\right)\right] \times \frac{I_{sij}}{Q_{sij}} \exp\left[-H_{sij}(\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2)^2\right] \exp\left[-\frac{1}{8w_s^2 Q_{sij}}(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2)^2\right] \\ \left. \times \exp\left[ikT_{sij}\left(\boldsymbol{\rho}_1^2 - \boldsymbol{\rho}_2^2\right)\right] \exp\left[-\frac{\boldsymbol{\rho}_1^2 + \boldsymbol{\rho}_2^2}{W^2}\right] \right\} \rho_1 \rho_2 d\rho_1 d\rho_2 d\varphi_1 d\varphi_2. \end{aligned} \tag{21.49}$$

Substituting the following relationship, $\boldsymbol{\rho}_1 \boldsymbol{\rho}_2 = \rho_1 \rho_2 \cos(\varphi_1 - \varphi_2)$, Eq. (21.49) can be arranged as

$$\begin{aligned} P_{IFij} = 2\Re^2 \frac{I_{oij} I_{sij}}{Q_{oij} Q_{sij}} \int_0^{2\pi} \int_0^{2\pi} \int_0^{\infty} \int_0^{\infty} \mathrm{Re}\left\{ \exp\left[-g_{ij}\left(\rho_1^2 + \rho_2^2\right)\right] \times \exp\left[ik\left(T_{sij} - T_{oij}\right)\left(\rho_1^2 - \rho_2^2\right)\right], \right. \\ \left. \times \exp\left[2q_{ij}\rho_1\rho_2\cos(\varphi_1 - \varphi_2)\right] \right\} \rho_1 \rho_2 d\rho_1 d\rho_2 d\varphi_1 d\varphi_2 \end{aligned} \tag{21.50}$$

where

$$\begin{aligned} g_{ij} &= H_{oij} + \frac{1}{8w_o^2 Q_{oij}} + H_{sij} + \frac{1}{8w_s^2 Q_{sij}} + \frac{1}{W^2} \\ q_{ij} &= H_{oij} - \frac{1}{8w_o^2 Q_{oij}} + H_{sij} - \frac{1}{8w_s^2 Q_{sij}}. \end{aligned} \tag{21.51}$$

Through a simplifying calculation, we can obtain

$$P_{IFij} = \frac{2\pi^2 \Re^2 I_{oij} I_{sij}}{Q_{oij} Q_{sij}\left(g_{ij}{}^2 + s_{ij}{}^2 - q_{ij}{}^2\right)}, (i, j = x, y). \tag{21.52}$$

The local optical power is

$$P_{oii} = \Re \iint J_{oii}(\boldsymbol{\rho}, \boldsymbol{\rho}) d^2\boldsymbol{\rho} = \frac{2\Re\pi I_{oii} W^2 w_o^2}{W^2 + 2w_o^2 Q_{oii}}, (i = x, y). \tag{21.53}$$

According to the above results,

$$SNR_{EGSM} = \frac{\sum_{i,j=x,y} P_{IFij}}{eB\sum_{i=x,y} P_{oii}}. \tag{21.54}$$

Similarly, the signal-light power is

$$P_{sii} = \Re \iint J_{sii}(\boldsymbol{\rho},\boldsymbol{\rho}) d^2\boldsymbol{\rho} = \frac{2\Re\pi I_{sii} W^2 w_s^2}{W^2 + 2w_s^2 Q_{sii}}, (i = x, y). \tag{21.55}$$

The heterodyne efficiency of the heterodyne detection can be expressed as

$$\eta_{EGSM} = \frac{\sum_{i,j=x,y} P_{IFij}}{2\sum_{i=x,y} P_{oii} \sum_{i=x,y} P_{sii}}. \tag{21.56}$$

According to Eqs. (21.54) and (21.56), $\mathrm{SNR_{EGSM}}$ can be written as

$$\begin{aligned} SNR_{EGSM} &= \frac{\sum_{i=x,y} P_{sii} \sum_{i,j=x,y} P_{IFij}}{eB\sum_{i=x,y} P_{oii} \sum_{i=x,y} P_{sii}} \\ &= \frac{2\sum_{i=x,y} P_{sii}}{eB} \eta_{EGSM}. \end{aligned} \tag{21.57}$$

The bit-error rate of the receiver is also an important parameter to measure the performance of a coherent optical communication system. The expression for the bit-error rate of a binary phase-shift keying (BPSK) receiver is [119]

$$BER = \frac{1}{2} erfc\left(\frac{SNR_{EGSM}}{2}\right)^{1/2}. \tag{21.58}$$

The detection system sensitivity refers to the minimum input signal power that can be detected under a certain bit-error rate. It can be obtained from Eq. (21.59):

$$\sum_{i=x,y} P_{sii} = \frac{eB}{2\eta_{EGSM}} SNR_{EGSM}, \tag{21.59}$$

where B is the bandwidth of the detector.

Experimental Study

The influence of an oblique transmission of an EGSM beam under non-Kolmogorov atmospheric turbulence on the sensitivity of differential detection under different polarization states was studied by numerical analysis.

As can be seen from Eq. (21.59), the sensitivity of the EGSM beam for heterodyne detection is determined by the turbulence size, turbulence scale, detector size, transmission distance, height of receiving end, height of light source H, etc. This study mainly analyzes how each factor affects the sensitivity of the heterodyne-detection system when the given bit-error rate $\mathrm{BER} = 10^{-9}$.

Firstly, the values of several global parameters are given. In the subsequent analysis, the values of these parameters are all given here, unless otherwise specified. The wavelength value for $\lambda = 650$ nm, internal and external turbulence scale values are $l_0 = 5 \times 10^{-3}$ m and $L_0 = 1$ m, power-spectrum power law of $\alpha = 3.7$, atmospheric structure constant $C_0 = 1.7 \times 10^{-14}$ m$^{-2/3}$, radius of the vibration signal light and sending light $w_o = w_s = 3$ mm, detector diameter $D = 3$ mm, quantum efficiency $\eta = 0.8$, Planck's constant $h = 6.623 \times 10^{-34}$ J·s, carrier frequency $v = 1.9 \times 10^{14}$ m/s, and detector bandwidth $B = 200$ MHz.

To ensure that the polarization characteristics of the laser remain unchanged in the atmospheric channel, the beam space lengths of the signal light and local light are taken as $\delta_{oxy} = \delta_{oyx} = 0.2$ mm, $\delta_{oxx} = \delta_{oyy} = 0.2$ mm, $\delta_{sxx} = \delta_{syy} = 0.4$ mm, and $\delta_{sxy} = \delta_{syx} = 0.4$ mm, respectively.

To analyze the influences of various factors on the sensitivity of heterodyne detection under different polarization states during the oblique transmission of an EGSM beam in a more detailed and visual manner, the sensitivity of the EGSM beam under the three states of complete polarization, unpolarization, and partial polarization for heterodyne detection was analyzed in detail in this study, as shown in Figs. 21.38, 21.39, and 21.40.

According to Eq. (21.48), when $I_{oxy} = I_{sxy} = 1$, $I_{oyx} = I_{syx} = 1$, $I_{oxx} = I_{oyy} = 1$, and $I_{sxx} = I_{syy} = 1$, the degree of polarization $P = 1$; then the signal light and the local

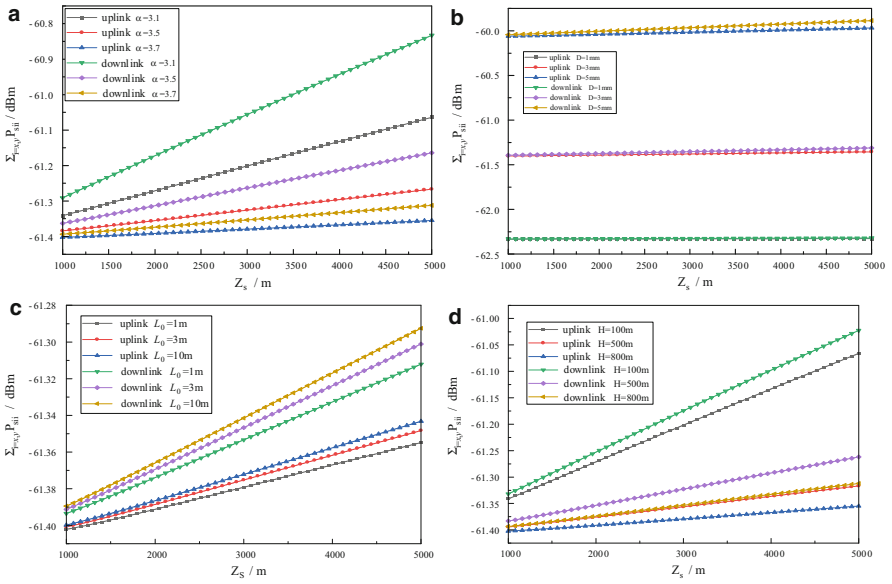


Fig. 21.38 System sensitivity of a fully polarized EGSM beam for heterodyne detection varying with the transmission distance, both up and down. (**a**) Different power-spectrum power laws α. (**b**) Different detector diameters D. (**c**) Different outer turbulence scales L_0. (**d**) Different heights of receiving end or light source H

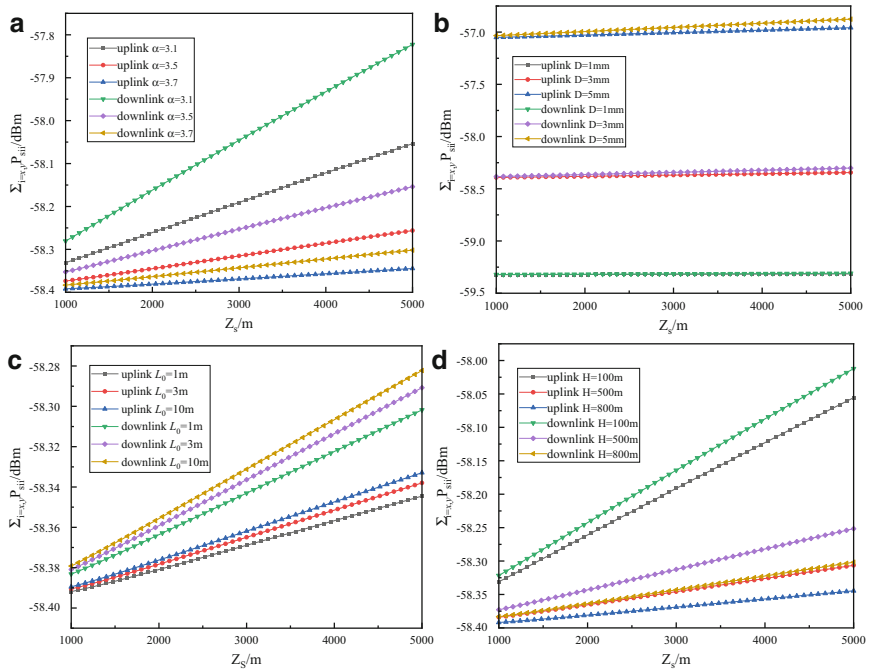


Fig. 21.39 System sensitivity of an unpolarized EGSM beam for heterodyne detection varying with the transmission distance, both up and down. (**a**) Different power-spectrum power laws α. (**b**) Different detector diameters D. (**c**) Different outer turbulence scales L_0. (**d**) Different heights of receiving end or light source H

oscillator light are fully polarized beams. Figure 21.38a–d respectively show the variation curves of the system sensitivity with the transmission distance when fully polarized EGSM beams are used for heterodyne detection under different power-spectrum power laws, detector diameters, outer turbulent scales, height of receiving end, or height H of light source, when the beam is transmitted up and down through atmospheric turbulence.

As can be seen from Fig. 21.38, the detection sensitivity of the system decreases with the increase in transmission distance. Because the beam is affected by the refractive index of the random fluctuation of the atmosphere when it is transmitted in atmospheric turbulence, the wavefront phase distortion will be more serious when the transmission distance is longer; hence, the turbulence will have a greater impact on the beam, resulting in a decrease in sensitivity.

Figure 21.38a shows that under a certain transmission distance, the detection sensitivity of the system increases with the increase of the power-spectrum power law. When the power-spectrum power law α is 3.1, the turbulence has a greater influence on the beam and the sensitivity is minimal. The sensitivity change of a turbulent downlink is greater than that of a turbulent uplink.

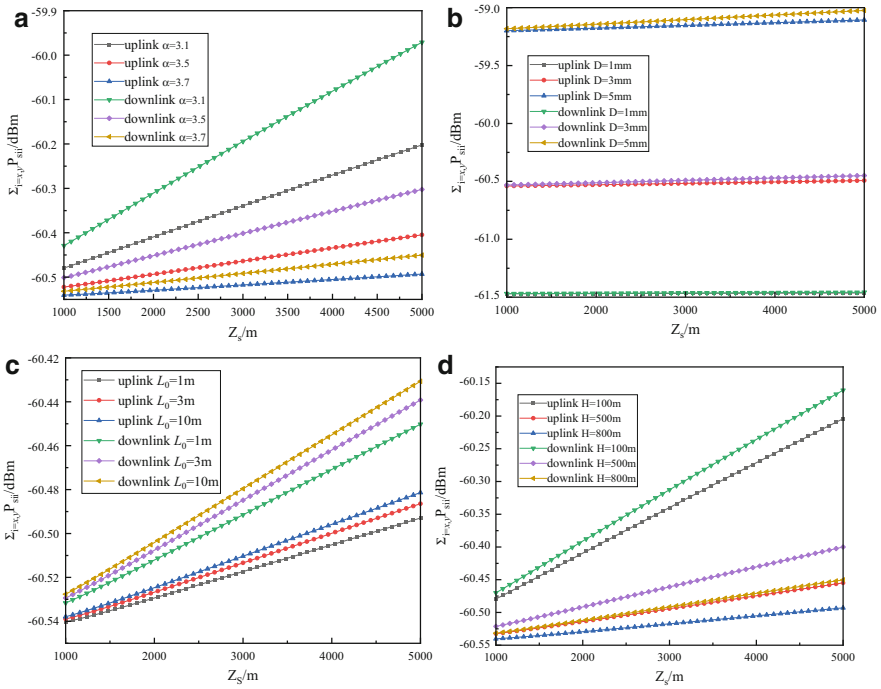


Fig. 21.40 System sensitivity of a partially polarized EGSM beam for heterodyne detection varying with the transmission distance, both up and down. (**a**) Different power-spectrum power laws α. (**b**) Different detector diameters D. (**c**) Different outer turbulence scales L_0. (**d**) Different heights of the receiving end or light source H

As can be seen from Fig. 21.38b, under a certain transmission distance, the detection sensitivity of the system decreases with the increase of the diameter of the detector. Considering the limitation of the spatial coherence of the signal light, the two beams with a smaller detector size have a better match, so the detection sensitivity of the system with a smaller detector size is larger. When the transmission distance increases gradually, the system sensitivity changes slowly with the transmission distance.

Figure 21.38c shows that, under a certain transmission distance, the detection sensitivity of the system decreases with the increase of the outer turbulence scale, because as the outer turbulence scale increases, the corresponding intensity of the atmospheric turbulence gradually increases, and the more serious the diffraction during beam transmission. This results in a decrease in the spatial coherence of the signal light and a decrease in sensitivity.

Figure 21.38d shows that under a certain transmission distance, the detection sensitivity of the system decreases with the decrease of the height of different receiving ends or the height H of the light source. In the transmission process, when the transmission height is less than 10 km, the refractive-index structure constant of atmospheric turbulence increases with the decrease of the height.

When the transmission distance remains unchanged, the smaller H is, the closer it is to a horizontal transmission, and the transmission path of the signal light in the strong turbulent region increases, resulting in a decrease of system sensitivity.

In addition, as can be seen from Fig. 21.38 as a whole, the sensitivity of the uplink is greater than that of the downlink, which is caused by the difference in the turbulence terms M_1, M_2, and M_3 in Eq. (21.39). In Eq. (21.39), the value of the turbulence term of the uplink is smaller than that of the turbulence term of the downlink. Therefore, the sensitivity is larger.

According to Eq. (21.48), when $I_{oxy} = I_{sxy} = 0$, $I_{oyx} = I_{syx} = 0$, $I_{oxx} = I_{oyy} = 1$, and $I_{sxx} = I_{syy} = 1$, the degree of polarization $P = 0$, then the signal light and local light are unpolarized beams. In Figs. 21.39a–d, when unpolarized EGSM beams are used for heterodyne detection, the sensitivity of the system varies with the transmission distance, when the beams are transmitted up and down through atmospheric turbulence under different power-spectrum power laws, detector diameters, outer turbulence scales, heights of the receiving end, or heights H of the light source.

As can be seen from the figure, with the increase in transmission distance, the smaller the power-spectrum power law, the larger the detector diameter, the larger the turbulence intensity, the lower the height of the receiving end or the height H of the light source, and the smaller the detection sensitivity of the system. By comparing Figs. 21.38 and 21.39, it can be seen that under the same link conditions, the sensitivity of an unpolarized EGSM beam for heterodyne detection is lower than that of a fully polarized beam, on the whole.

According to Eq. (21.48), when $I_{oxy} = I_{sxy} = 0.8$, $I_{oyx} = I_{syx} = 0.8$, $I_{oxx} = I_{oyy} = 1$, and $I_{sxx} = I_{syy} = 1$, the degree of polarization $P = 0.8$; then, the signal light and local light are partially polarized beams. In Fig. 21.40a–d, partially polarized EGSM beams are used for heterodyne detection. With different power-spectrum power laws, detector diameters, outer turbulence scales, receiving end heights or light-source heights H, the system sensitivity varies with the transmission distance when the beam is transmitted up and down through atmospheric turbulence.

By comparing Figs. 21.38, 21.39, and 21.40, it can be seen that under the same link conditions, the variation trend of the different polarization states is basically similar. The difference is that the detection sensitivity of a partially polarized EGSM beam for heterodyne detection is between that of an unpolarized beam and that of a fully polarized beam.

Because partially polarized light can be regarded as a combination of fully polarized light and unpolarized light, the performance of the system should be between unpolarized light and fully polarized light, when a partially polarized EGSM beam is used for heterodyne detection. It was obtained by numerical analysis that the optimal state sensitivity of a partially polarized EGSM beam for heterodyne detection was −61.5 dBm, which is between the optimal state of a fully polarized beam, −62.3 dBm, and the optimal state of unpolarized light, −59.3 dBm.

Based on the non-Kolmogorov turbulence spectrum, a theoretical model of the sensitivity of an EGSM beam for the heterodyne detection of an oblique transmission was established. Based on the numerical results, the system performance of EGSM beams with different polarization states, when transmitting up and down in

atmospheric turbulence, was studied in detail. The results showed that the sensitivity of the system was restricted by the power-spectrum power law α, and the detection sensitivity of fully polarized, partially polarized, and unpolarized EGSM beams for heterodyne detection increased with the increase of power-spectrum power law α.

Under the same conditions, the sensitivity of the EGSM beam was greater when it was transmitted upward in atmospheric turbulence than when it was transmitted downward. The detection sensitivity of an EGSM beam in the partially polarized state was between the fully polarized state and unpolarized state, and the sensitivity in the fully polarized state was the highest.

The sensitivity of a heterodyne-detection system can be adjusted by controlling the detector parameters, local-vibration parameters, transmission distance, turbulence scale, receiving height, etc. For a certain detector, a higher detection sensitivity can be obtained by adjusting the optimal value of the parameters of indirect factors, in addition to detector parameters.

Effect of the Optical Spatial Distribution on the Mixing Efficiency

Coherent optical communication systems are widely used in inter-satellite communication because of their high communication rate, good confidentiality, long-distance information transmission, and the small impact of background light [1]. Compared with intensity modulation/direct detection (IM/DD), coherent detection has high sensitivity and approaches the quantum detection limit [2]; however, the requirements of the receiving optical system are stricter, which not only requires a higher local power, but also ensures the efficient mixing of signal light and local light.

The mixing efficiency can be improved by ensuring as much as possible coherence matching between the signal light and local light. Coherence matching includes spatial-distribution (including amplitude and phase) and polarization-state matching. Polarization-state matching between the signal light and local vibration light can enable the beam to be realized through specific optical components, etc.

Scholars have more intently observed the influence of the spatial distribution of light on the mixing efficiency [120]. David et al. [121] pointed out that the optimal local light distribution should be the same as the signal-field distribution on the detector surface. On a circular detector, the signal light was the optimal detector radius when the local light was uniformly distributed. Cohen et al. studied the influence of lost supporting characters and the spot radius on mixing efficiency. It was found that the mixing efficiency was insensitive to the spatial-angle mismatch when the signal light had an Airy distribution.

Tanaka et al. [122] studied the changes in mixing efficiency of the system when the incident direction of the signal light was tilted and offset, under the premise that both the signal light and local light had a Gaussian distribution. Xiang et al. [123] studied the optical heterodyne efficiency of the combination of an Airy signal light

and Gaussian local oscillator light, and analyzed the optical mixing-efficiency loss caused by different aberrations, such as tilt, defocus, spherical aberration, coma, and astigmatism.

Liu et al. [124] studied the efficiency of mixing a signal light with a Gaussian and uniform distribution and different local modes in an interplanetary coherent optical communication system. Li et al. [125] studied the influence of collimation loss on the mixing efficiency of external error detection in a heterodyne-detection system, where the signal light had an Airy distribution and the local oscillator light had a Gaussian distribution.

Coherent Detection System

The schematic diagram of a coherent-detection system is shown in Fig. 21.41. After the signal light is transmitted through the atmospheric channel and reaches the receiving end, it synthesizes a beam with the local oscillator light emitted by the local oscillator laser, and then converges to the surface of the photodetector through the focusing lens for frequency mixing [126, 127].

The essence of frequency mixing is the interference between the signal light and local light on the surface of the photodetector. When two light beams are superimposed, the phase difference is fixed and the polarization direction is the same, which are the necessary conditions for the interference of two beams [128, 129]. The output IF signal of the photodetector contains multiple types of noise, so the mixing-efficiency evaluation index is introduced to evaluate the mixing performance of the system [130].

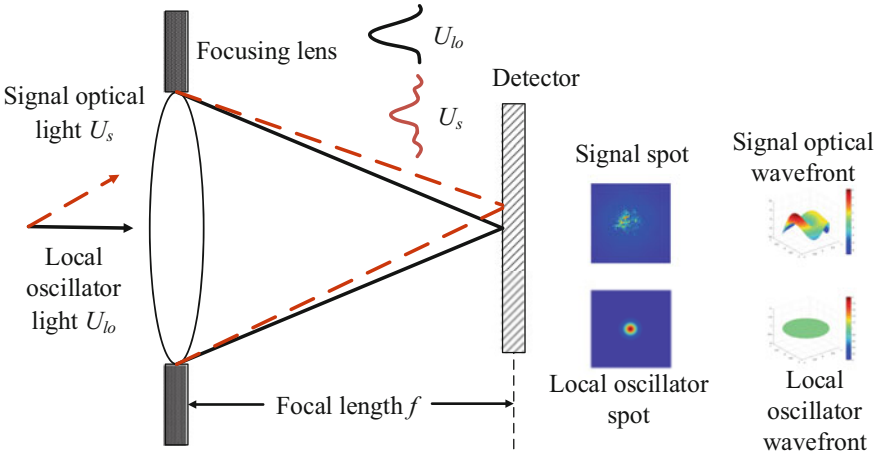


Fig. 21.41 Schematic diagram of a coherent-detection system

Mixing Efficiency of Coherent Detection Under Ideal Conditions

The mixing efficiency of optical heterodyne detection is the ratio of the IF signal power obtained from actually mixing two optical fields to the IF signal power obtained from the ideal mixing of two optical waves [131]. The ideal case here means that the spatial distribution of the two beams is the same, including amplitude and phase.

Assuming that the geometric centers of the two light fields coincide and the detector is distributed in circular symmetry, the mixing efficiency can be expressed as

$$\eta = \frac{\left|\int_0^{2\pi}\int_0^{r_0} U_s(r,\varphi)U_{lo}(r,\varphi)\exp[i\Delta\varphi(r,\varphi)]rdrd\varphi\right|^2}{\int_0^{2\pi}\int_0^{r_0}[U_{lo}(r,\varphi)]^2 rdrd\varphi\int_0^{2\pi}\int_0^{\infty}[U_s(r,\varphi)]^2 rdrd\varphi}, \tag{21.60}$$

where r_0 is the probe radius, $U_s(r, \varphi)$ is the detector signal-light amplitude distribution, $U_{lo}(r, \varphi)$ is the probe on the surface of the vibration amplitude of the light distribution, and $\Delta\varphi(r, \varphi)$ is the phase difference of the local vibration light and the signal light.

When the phase effect is not considered and the geometric centers of the signal light and local vibration light are assumed to coincide without an alignment error, the mixing efficiency Eq. (21.60) can be simplified as [132]

$$\eta = \frac{\left|\int_0^{r_0} U_s(r)U_{lo}(r)rdr\right|^2}{\int_0^{r_0}[U_{lo}(r)]^2 rdr\int_0^{\infty}[U_{lo}(r)]^2 rdr}. \tag{21.61}$$

Mixing Efficiency of Coherent Detection in a Turbulent Flow

The above discussion is based on the fact that the phase difference between the signal light and the local oscillator light is constant. However, in practice, when the signal light passes through the atmosphere, the uneven refractive index of the atmosphere will distort the phase of the signal light wavefront [133]. The additional phase difference brought by atmospheric turbulence will decrease the mixing efficiency and affect the performance of the coherent optical communication system.

The expression of a signal light transmitted through atmospheric turbulence is

$$E_{S1}(r,\varphi,t) = U_{S1}(r,\varphi)\exp[i(\omega_{s1}t+\varphi_{s1})] = U_S(r,\varphi,t)\exp(\Phi), \tag{21.62}$$

where U_{s1}, ω_{s1}, and φ_{s1}, respectively, refer to the amplitude, angular frequency, and initial phase of the signal light after atmospheric turbulence; Φ is the expression for atmospheric turbulence caused by amplitude and phase fluctuations. $U_s(r, \varphi, t)$ is the

probe at any point on the complex amplitude of the signal light field. The relation of $U_s(r, \varphi, t) = U_s(r, \varphi)\exp[i(\omega_0 t + \varphi_s)]$ changes over time, where U_s, ω_s, and φ_s are, respectively, the amplitude, angular frequency, and initial phase of the signal light. The $U_s(r, \varphi, t)$ expression in Eq. (21.62) is [18]

$$\Phi = \ln \frac{U_{S1}(r, \varphi)}{U_S(r, \varphi)} + i[\Phi_{S1} - \Phi_S] = \chi + iS, \tag{21.63}$$

where χ is the light-intensity change, S is the wavefront phase distortion, ln is the natural logarithm function of e, and Φ_S and Φ_{S1} are, respectively, the signal light affected by atmospheric turbulence before and after the joint expression of amplitude and phase fluctuations. We substitute Eq. (21.63) into Eq. (21.62) to obtain

$$U_{S1} = U_S(r, \varphi) \exp[i(\omega_S t + \varphi_S)] \exp(\chi + iS). \tag{21.64}$$

Equation (21.64) is substituted into Eq. (21.63) as the signal light:

$$\eta = \frac{\left| \int_0^{r_0} U_s(r) U_{lo}(r) \exp(-iS) r dr \right|^2}{\int_0^{r_0} [U_{lo}(r)]^2 r dr \int_0^{\infty} [U_{lo}(r)]^2 r dr}, \tag{21.65}$$

where S is the wavefront phase distortion, and $U_s(r)$ and $U_{lo}(r)$ are the amplitude distributions of the signal light and local light on the detector surface, respectively.

From Eq. (21.65), it can be seen that the mixing efficiency of coherent detection is affected by atmospheric turbulence. Because a Gaussian signal light is transmitted over a long distance, the optical-field distribution when it reaches the detector surface at the receiving end is a plane wave [134]. Phase fluctuations under the influence of atmospheric turbulence can be described by Kolmogorov statistical theory, and phase fluctuations of plane waves can be expressed as [135]

$$S = 2.914 C_n^2 L k^2 \rho^{5/3}. \tag{21.66}$$

In Eq. (21.66), C_n^2 is the atmospheric refractive-index structure constant, L is the signal optical-transmission distance, k is the wavenumber, and r is the length of the two-dimensional space vector between the inner and outer turbulence scales. By substituting Eq. (21.66) into Eq. (21.65), the mixing efficiency of a coherent optical communication system affected by atmospheric turbulence can be written as

$$\eta = \frac{\left| \int_0^{r_0} U_s(r) U_{lo}(r) \exp\left(-i\left(2.914 C_n^2 L k^2 \rho^{5/3}\right)\right) r dr \right|^2}{\int_0^{r_0} [U_{lo}(r)]^2 r dr \int_0^{\infty} [U_{lo}(r)]^2 r dr}. \tag{21.67}$$

A Gaussian beam is typically used as a signal light in coherent optical communication systems. The light spot of a Gaussian beam becomes larger after a long-distance transmission, and the signal light received by the detector takes up a small proportion, which can be approximately considered to be evenly distributed [136]. Selecting the local oscillator light that best matches the signal light is helpful to improve the mixing efficiency of the system.

The following will be divided into two categories, according to whether the signal light is Gaussian, uniform, or a vortex beam, and the mixing efficiency when the local light is a different spatial distribution will be discussed. In the subsequent analysis process, the wavelength λ and radius of the Gaussian beam ω_0 are unified as $\lambda = 1550$ nm and $\omega_0 = 1.6\,\lambda$.

Mixing Efficiency of a Gaussian-Distributed Signal Light and Various Local Vibration Lights

After a lens-focused Gaussian beam reaches the surface of the detector, it still has a Gaussian distribution; however, the coefficient of the Gaussian distribution function changes, because in the process of computing the mixing efficiency, the molecular denominator will appear at the same time. The constant coefficient of the vibration light and signal light expressions can be reduced by a fraction.

We set the expression of a Gaussian signal light as $U_s(r) = \exp(-r^2/{\omega_0}^2)$. The local vibration light can be Gaussian or Airy: $U_s(r) = 2\,J_1(\pi r/F\lambda)\,/\,(\pi r/F\lambda)$. In the uniform distribution, $U_s(r) = 1$, J_1 is the first type of first-order Bessel function, and F is the F number of the receiving optical system. In the below analysis, we set $X = \pi r/F\lambda$, $X_0 = \pi r_0/F\lambda$, $Q = r/\omega$ and $Q_0 = r/\omega_0$.

Local Vibration Light Is Uniformly Distributed

When the local vibration light has a uniform distribution, it can be concluded that under the influence of atmospheric turbulence, the mixing efficiency η_{G-U} is as follows:

$$\eta_{G-U} = \frac{\left|\int_0^{r_0} \exp\left(-Q^2\right)\exp\left(-i\left(2.914C_n^2Lk^2\rho^{5/3}\right)\right)rdr\right|^2}{\int_0^{r_0} rdr\int_0^{\infty}\left[\exp\left(-Q^2\right)\right]^2 rdr}. \tag{21.68}$$

Figure 21.42 shows the mixing-efficiency curves when the transmission distance $L = 500$ m and two-dimensional space radius-vector length is $\rho = 0.1$ m, for different values of turbulence (C_n^2). The mixing efficiency has a maximum value of 81.45%, with the increase of the parameters of the receiving optical system. It decreases rapidly with the increase of the parameters of the receiving optical system, indicating that the mixing efficiency of this model is more sensitive to these

parameters; therefore, it is necessary to adjust these parameters to obtain a higher mixing efficiency. The mixing efficiency decreases with the increase of C_n^2, and the mixing efficiency is less than 40% when $C_n^2 = 10^{-14}\ \mathrm{m}^{-2/3}$.

Local Vibration Light Has a Gaussian Distribution

When the local vibration light has a Gaussian distribution, it can be concluded that the mixing efficiency η_{G-G} under the influence of atmospheric turbulence is

$$\eta_{G-G} = \frac{\left|\int_0^{r_0} \exp\left(-Q^2\right) \exp\left(-i\left(2.914 C_n^2 L k^2 \rho^{5/3}\right)\right) rdr\right|^2}{\int_0^{r_0} \left[\exp\left(-Q^2\right)\right]^2 rdr \int_0^{\infty} \left[\exp\left(-Q^2\right)\right]^2 rdr}. \tag{21.69}$$

In combination with Eq. (21.69), the variation curves of the mixing efficiency with different values of C_n^2 under turbulent conditions are obtained, as shown in Fig. 21.43. Without turbulence, the mixing efficiency rapidly increases to the maximum value of 100% with the increase of the parameters of the receiving optical system. The mixing efficiency gradually decreases under the influence of atmospheric turbulence. When $C_n^2 = 10^{-16}\ \mathrm{m}^{-2/3}$, the mixing efficiency is greater than 80%, indicating that the mixing performance of this model is less affected by atmospheric turbulence.

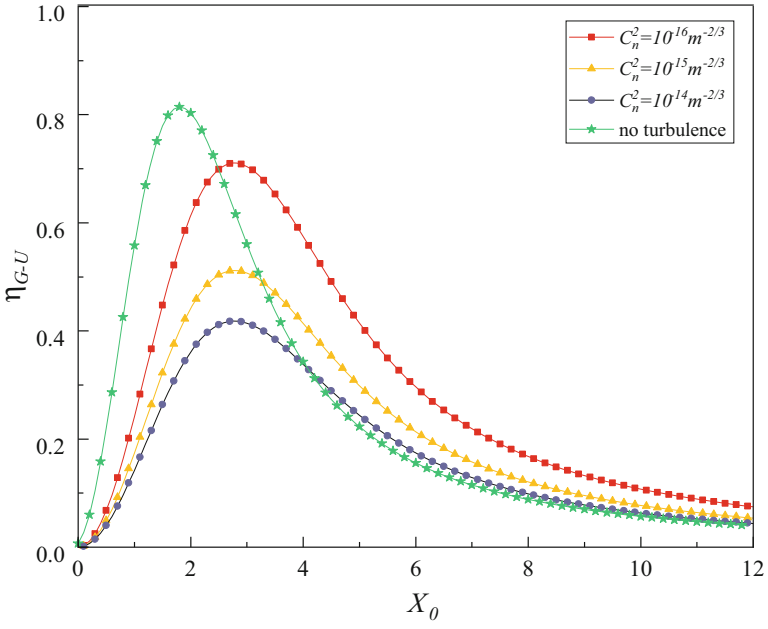


Fig. 21.42 Mixing efficiency of a Gauss–uniform distribution model

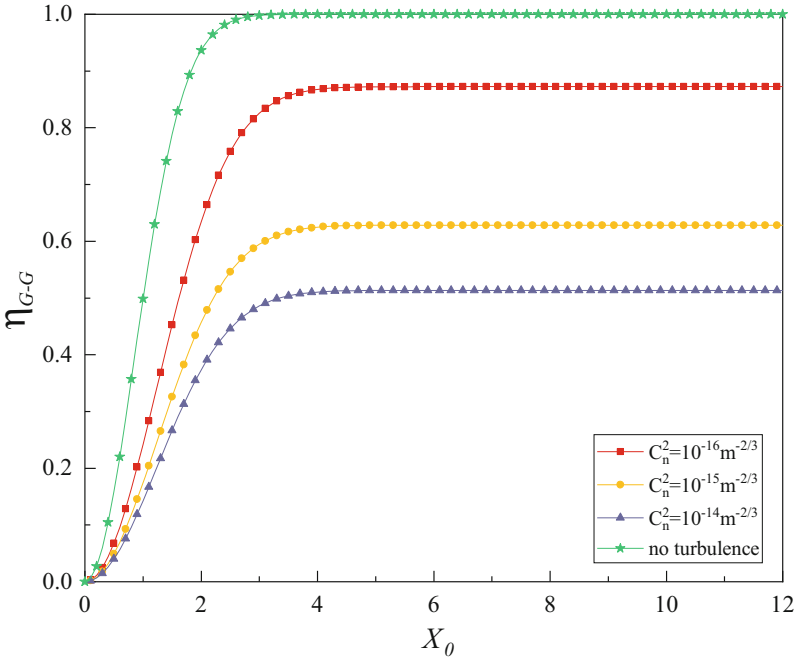


Fig. 21.43 Mixing efficiency of a Gauss–Gaussian distribution model

Local Oscillator Light Has an Airy Distribution

When the local oscillator light has an Airy distribution, it can be concluded that the mixing efficiency η_{G-A} under the influence of atmospheric turbulence is as follows:

$$\eta_{G-A} = \frac{\left|\int_0^{r_0} [J_1(X)/(X)] \exp\left(-Q^2\right) \exp\left(-i\left(2.914 C_n^2 L k^2 \rho^{5/3}\right)\right) rdr\right|^2}{\int_0^{r_0} [J_1(X)/(X)]^2 rdr \int_0^{\infty} \left[\exp\left(-Q^2\right)\right]^2 rdr}. \tag{21.70}$$

Figure 21.44 shows the variation curve of the mixing efficiency when C_n^2 takes different values, under ideal conditions and turbulent conditions. In the case with no turbulence, the mixing efficiency can be close to 100% when appropriate parameters are selected for the receiving optical system. However, when the parameters of the receiving optical system continue to increase, the mixing efficiency decreases instead of increasing, but still remains at about 80%. With the increase of the turbulence intensity, the mixing efficiency gradually decreases. When C_n^2 is 10^{-14} $m^{-2/3}$, the mixing efficiency is less than 50%, which is 50% lower than that without turbulence.

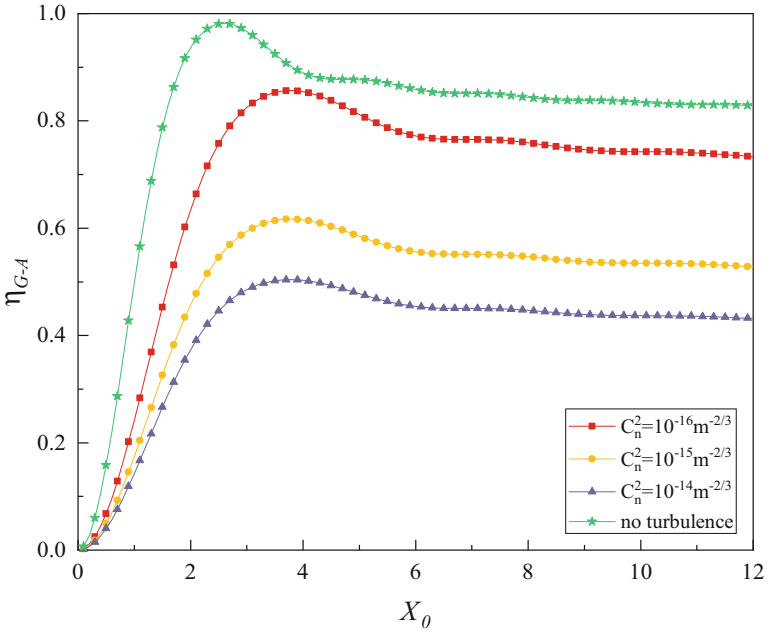


Fig. 21.44 Mixing efficiency of a Gauss–Airy distribution model

Mixing Efficiency When the Signal Light Has an Airy Distribution with Varying Local Vibration Lights

When a uniform signal light passes through the diffraction of the lens in Fig. 21.41, the spatial distribution becomes an Airy distribution when it focuses on the detector surface. The signal light is considered to have an Airy distribution below [19]. The local oscillator light can be a Gaussian, Airy, or uniformly distributed beam, and the mixing efficiency can be calculated by various equations.

Local Vibration Light Is Uniformly Distributed

When the local vibration light has a uniform distribution, it can be concluded that, under the influence of atmospheric turbulence, the mixing efficiency η_{A-U} is as follows:

$$\eta_{A-U} = \frac{\left|\int_0^{r_0} [J_1(X)/(X)] \exp\left(-i\left(2.914 C_n^2 L k^2 \rho^{5/3}\right)\right) rdr\right|^2}{\int_0^{r_0} rdr \int_0^{\infty} [J_1(X)/(X)]^2 rdr}. \tag{21.71}$$

According to Eq. (21.71), the variation curve of the mixing efficiency when C_n^2 takes different values is shown in Fig. 21.45. When the signal light has an Airy

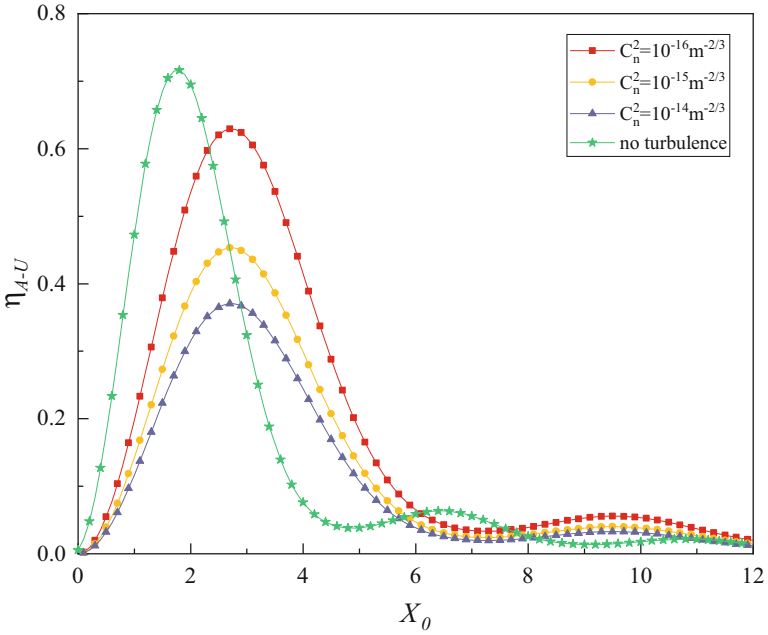


Fig. 21.45 Mixing efficiency of an Airy–uniform distribution model

distribution and the local oscillator light is uniformly distributed, the mixing efficiency quickly reaches the maximum value of 72.1% and then decreases. This indicates that this model is greatly affected by the parameters of the receiving optical system, and the conditions for reaching the maximum value are strict. Moreover, when atmospheric turbulence exists, the parameters of the receiving optical system required to reach the maximum mixing efficiency increase. When $C_n^2 = 10^{-14}\ \mathrm{m}^{-2/3}$, the mixing efficiency is less than 40%.

Local Oscillator Light Has an Airy Distribution

When the local oscillator light has an Airy distribution, it can be concluded that the mixing efficiency $\eta_{A\text{-}A}$ under the influence of atmospheric turbulence is as follows:

$$\eta_{A-A} = \frac{\left|\int_0^{r_0} [J_1(X)/(X)]^2 \exp\left(-i\left(2.914 C_n^2 L k^2 \rho^{5/3}\right)\right) rdr\right|^2}{\int_0^{r_0} [J_1(X)/(X)]^2 rdr \int_0^{\infty} [J_1(X)/(X)]^2 rdr}. \tag{21.72}$$

Combined with Eq. (21.72), the variation curve of the mixing efficiency when C_n^2 takes different values is shown in Fig. 21.46. When both the signal light and local light are Airy distributed, the mixing efficiency increases rapidly to about 80%

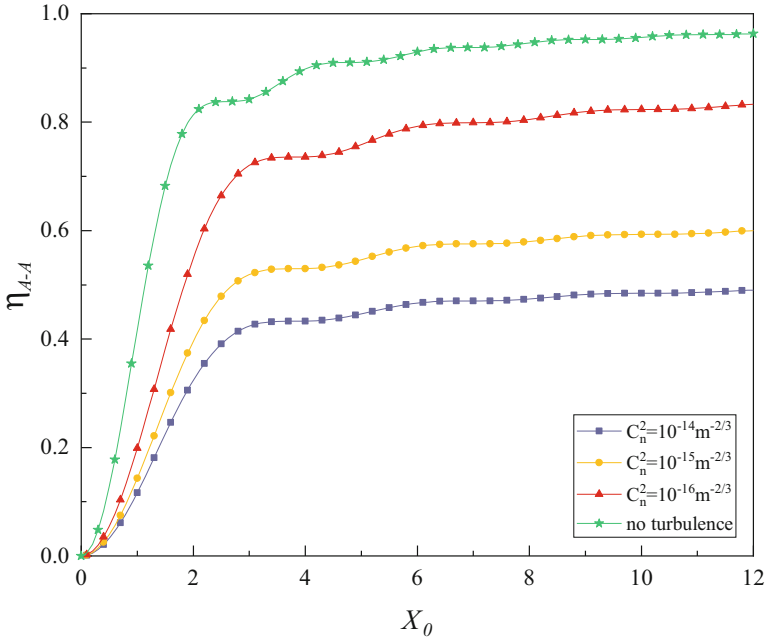


Fig. 21.46 Mixing efficiency of an Airy–Airy distribution model

with the increase of the parameters of the receiving optical system, then increases slowly, and gradually tends to 100%. The mixing efficiency decreases with the increase of C_n^2.

Local Light Has a Gaussian Distribution

When the local vibration light has a Gaussian distribution, it can be concluded that the mixing efficiency η_{A-G} under the influence of atmospheric turbulence is as follows:

$$\eta_{A-G} = \frac{\left|\int_0^{r_0} [J_1(X)/(X)] \exp\left(-Q^2\right) \exp\left(-i\left(2.914 C_n^2 L k^2 \rho^{5/3}\right)\right) rdr\right|^2}{\int_0^{\infty} [J_1(X)/(X)]^2 rdr \int_0^{r_0} \left[\exp\left(-Q^2\right)\right]^2 rdr}. \tag{21.73}$$

According to Eq. (21.73), the variation curve of the mixing efficiency when C_n^2 takes different values is shown in Fig. 21.47. When the signal light has an Airy distribution and the local vibration light has a Gaussian distribution, the mixing efficiency can reach 83.42% under the most suitable parameters of the receiving optical system. As can be seen from the figure, with the increase of the detection radius, the mixing efficiency linearly increases and reaches the maximum value of 83.42%. It then remains stable, indicating that the model has a strong adaptability to

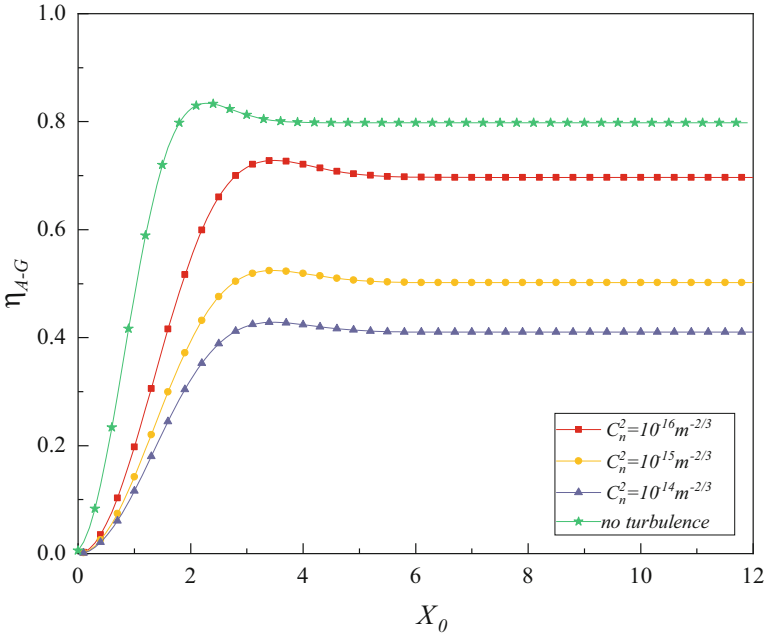


Fig. 21.47 Mixing efficiency of an Airy–Gauss distribution model

the receiving optical system. Under the influence of atmospheric turbulence, the mixing efficiency decreases with the increase of C_n^2.

The signal light has an Airy distribution and the local light has a Gaussian distribution, which is the most commonly used model in practical systems. The factors affecting the mixing efficiency of this model are analyzed in detail below. It can be seen from Eq. (21.73) that in the case of no turbulence, the mixing efficiency is related to the F number of the receiving optical system and the waist radius of the Gaussian beam. The two influencing factors are studied below. Figure 21.48 shows the corresponding mixing efficiencies for different X_0 and ω_0.

It can be seen from Fig. 21.48 that the mixing efficiency in the initial stage increases linearly with the increase of the parameters of the receiving optical system. This is because, with the increase of these parameters, the energy of the received signal light increases, so the mixing efficiency also increases. However, when the detection radius continues to increase, the mixing efficiency remains unchanged or even decreases. When the parameters of the receiving optical system increase to a certain extent, the background noise entering the detector also increases [128], which reduces the mixing efficiency. The mixing efficiency cannot continue to increase with the increase of the parameters of the receiving optical system.

Figure 21.49 shows the relation curve between the mixing efficiency and the waist radius of the Gaussian beam when the detector radius r_0 is 40 μm and the F number of the receiving optical system varies. It can be seen from Fig. 21.49 that the mixing efficiency of the system can be improved by selecting an appropriate

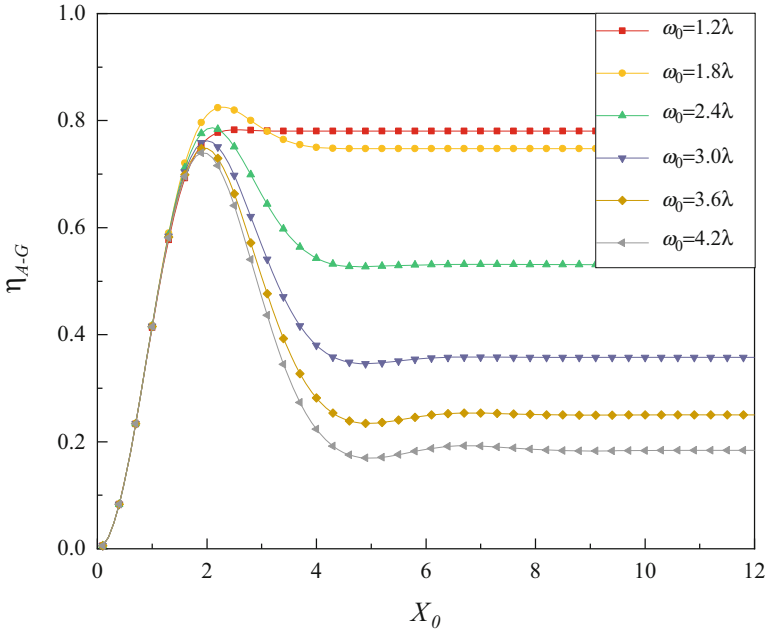


Fig. 21.48 Mixing efficiency of an Airy–Gaussian model with different ω_0

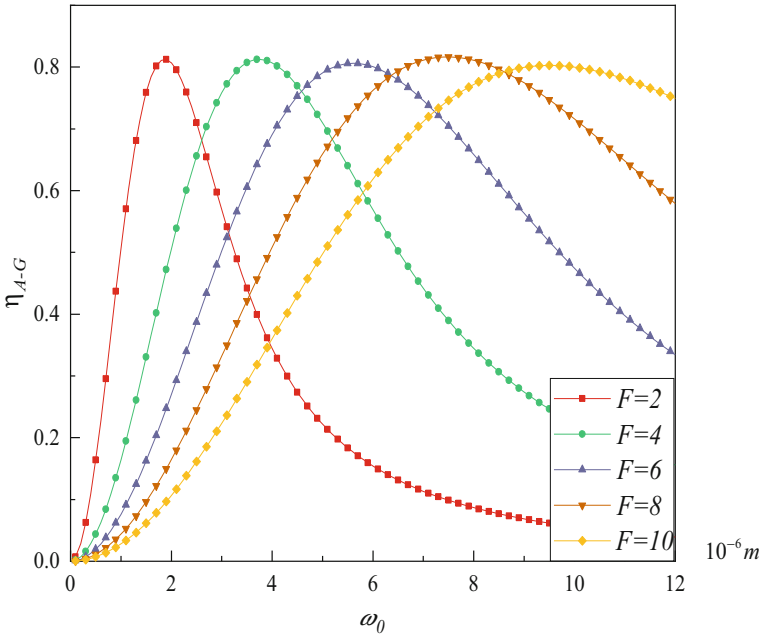


Fig. 21.49 Mixing efficiency of an Airy–Gaussian light, showing the relationship between the F number and ω_0

F value. When the selected F number is appropriate, the mixing efficiency can reach an ideal maximum. According to the trend of the curve, the larger the beam waist radius, the larger the F number required to achieve the ideal maximum.

In the following, the change in mixing efficiency caused by different receiving antennas is analyzed. A Zernike polynomial is used to represent the wavefront of the transmitted signal light. The Zernike polynomial containing the primary difference is [127]

$$\begin{cases} W_{\text{tilt}}(\rho,\varphi) = W_{111}\rho\cos\varphi \\ W_{\text{defocus}}(\rho,\varphi) = W_{020}\varphi^2 \\ W_{\text{spherical aberration}}(\rho,\varphi) = W_{040}\rho^4 \\ W_{\text{coma}}(\rho,\varphi) = W_{131}\rho^3\cos\varphi \\ W_{\text{astigmatism}}(\rho,\varphi) = W_{222}\rho^2\cos\varphi \end{cases}. \tag{21.74}$$

The above numerical shows the deviation of the actual wavefront from the ideal wavefront in the direction of the propagation distance. In Eq. 21.74, ρ is a normalized radius, i.e., $\rho = \rho/\mathrm{P}$; $\rho = 100$ μm for the radius of the receiving antenna. The equation for the mixing efficiency is as follows:

$$\eta = \frac{\left|\int_0^{2\pi}\int_0^{r_0} U_s(r,\varphi)U_{lo}(r,\varphi)\exp[ikW_X(\rho,\varphi)]rdrd\varphi\right|^2}{\int_0^{2\pi}\int_0^{r_0}[U_{lo}(r,\varphi)]^2rdrd\varphi\int_0^{2\pi}\int_0^{\infty}[U_s(r,\varphi)]^2rdrd\varphi}, \tag{21.75}$$

where $k = 2\pi/\lambda$, $kW_X(\rho, \varphi)$ is the antenna aberration introduced by the phase difference between the local vibration light and signal light, and W is the size of the wavefront aberration value.

As shown in Fig. 21.50, the primary aberration reduces the mixing efficiency to varying degrees, and the tilt difference has the most obvious effect on the mixing efficiency. When the aberration value is 0.2 λ, the mixing efficiency decreases by about 10%.

Mixing Efficiency When the Signal Light Is a Vortex Beam with Varying Local Vibration Lights

A Laguerre–Gaussian beam (LG) is a common vortex beam in optical communication systems. Gaussian beams are often used as the medium to detect vortex light in communication systems [137, 138]. Owing to the particularity of the vortex-beam phase, the mixing efficiency of vortex heterodyne detection will be studied from the angles of amplitude and phase matching. The feasibility of optical heterodyne detection for vortex optical communication systems is discussed. The mixing-

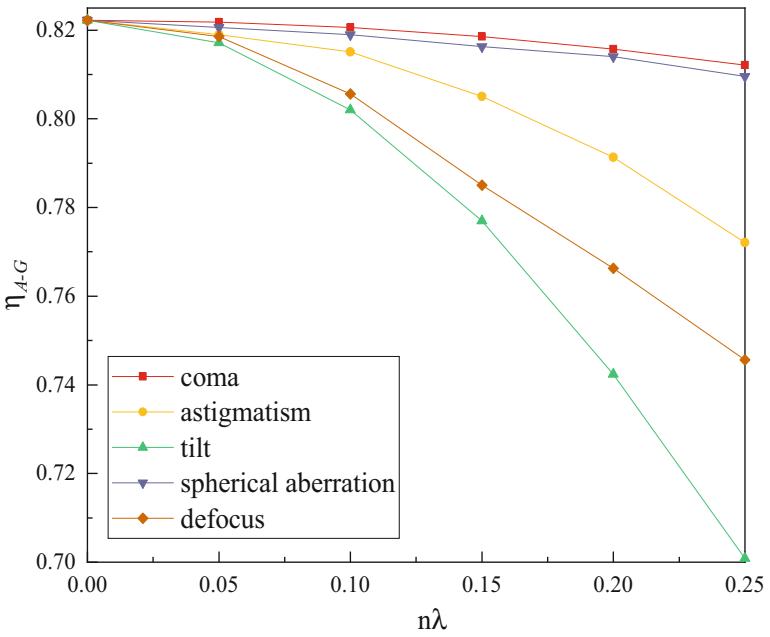


Fig. 21.50 Relationship between antenna aberrations and the mixing efficiency

efficiency changes are obtained when the signal light is an LG beam and the local light has Gaussian, Airy, and uniform distributions, and is a vortex light.

When the radial index is 0, the normalized amplitude of the LG beam carrying orbital angular momentum (OAM) at the transmission distance of 0 can be expressed as [139, 140]

$$U(r,\theta) = \sqrt{\frac{2}{\pi}}\frac{1}{\omega_0}\exp\left(-\frac{r^2}{\omega_0^2}\right)\sqrt{\frac{1}{|l!|}}\left[\frac{r\sqrt{2}}{\omega_0}\right]^{|l|}\exp(il\theta), \tag{21.76}$$

where l is the topological charge number, $\exp(il\theta)$ is the spiral phase factor, and ω_0 is the beam waist radius.

Local Light Has a Gaussian Distribution

When the local vibration light has a Gaussian distribution, assuming that the two phases match, we study the influence of the amplitude on the mixing efficiency, η_{G-L}:

$$\eta_{G-L} = \frac{\left|\int_0^{r_0}\sqrt{\frac{2}{\pi}}\frac{1}{\omega_0}\exp\left(-Q^2\right)\sqrt{\frac{1}{|l!|}}\left[\sqrt{2}Q\right]^{|l|}\exp\left(-Q^2\right)rdr\right|^2}{\int_0^{\infty}\left[\sqrt{\frac{2}{\pi}}\frac{1}{\omega_0}\exp\left(-Q^2\right)\sqrt{\frac{1}{|l!|}}\left[\sqrt{2}Q\right]^{|l|}\right]^2 rdr\int_0^{r_0}\left[\exp\left(-Q^2\right)\right]^2 rdr}, \tag{21.77}$$

where l is the topological charge number for LG beams and ω_0 is the Gaussian beam waist radius.

When the local oscillator light has a Gaussian distribution and the signal light is an LG beam, the relation between the mixing efficiency and the parameters of the receiving optical system under different topological charges is obtained according to Eq. (21.77), and the topological charges are set as 1, 3, and 5. The resulting change curve is shown in Fig. 21.51.

As can be seen from the figure, when the topological charge is 1 pentad, the mixing efficiency can reach about 80%; when the topological charge is 3, the mixing efficiency is only about 30%, and when the topological charge is 5, the mixing efficiency is only 10%. The mixing efficiency gradually decreases with the increase of the topological charge, and the parameters of the receiving optical system required by the mixing efficiency to stabilize gradually increase. After reaching the maximum, the mixing efficiency remains stable and does not change with the increase of the parameters of the receiving optical system.

According to Eq. (21.77), the mixing efficiency changes with the topological charge of the LG beam, as shown in Fig. 21.52. The mixing efficiency decreases linearly with the increase of the topological charge, and decreases to 50% when the topological charge is 3.

Assuming that the amplitudes of the signal light and local light match, the influence of the phase on the mixing efficiency is discussed. The mixing efficiency expression calculated by Eq. (21.77) is used, and the parameters of the receiving

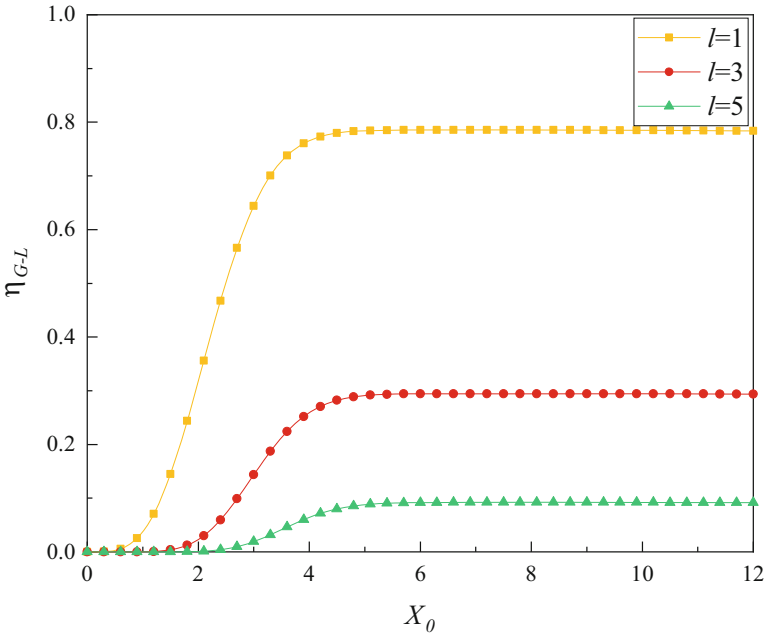


Fig. 21.51 Mixing efficiency of an LG beam when the local light has a Gaussian distribution

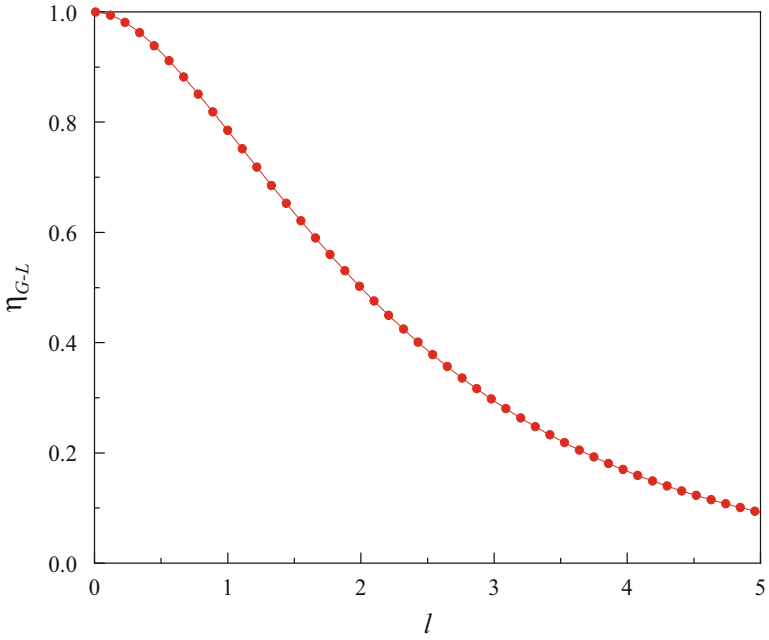


Fig. 21.52 Mixing efficiency changes with the topology load

optical system are set as fixed values. The relationship between the topological charge and mixing efficiency is studied, as shown in Fig. 21.53.

When the topological charge is greater than 1, the mixing efficiency is less than 10%. When the topological charge is an integer, the mixing efficiency is almost 0, and the mixing efficiency between two integer topological charges first increases and then decreases. In practical applications, the topological charges are mostly integers, so from the point of view of phase matching, Gaussian beams and LG beams cannot be mixed.

Local Oscillator Light Has an Airy Distribution

When the local light has an Airy distribution, assuming that the phases match, we investigate the relationship between the amplitude distribution and the mixing efficiency. The equation for the mixing efficiency η_{A-L} is as follows:

$$\eta_{A-L} = \frac{\left| \int_0^{r_0} \sqrt{\frac{2}{\pi}} \frac{1}{\omega_0} \exp\left(-Q^2\right) \sqrt{\frac{1}{|l|!}} \left[\sqrt{2}Q\right]^{|l|} \left[\frac{J_1(X)}{(X)}\right] rdr \right|^2}{\int_0^{\infty} \left[\sqrt{\frac{2}{\pi}} \frac{1}{\omega_0} \exp\left(-Q^2\right) \sqrt{\frac{1}{|l|!}} \left[\sqrt{2}Q\right]^{|l|} \right]^2 rdr \int_0^{r_0} \left[\left[\frac{J_1(X)}{(X)}\right] \right]^2 rdr}. \tag{21.78}$$

The relationship between the mixing efficiency and the parameters of the receiving optical system is calculated according to Eq. (21.78), and the topological charge

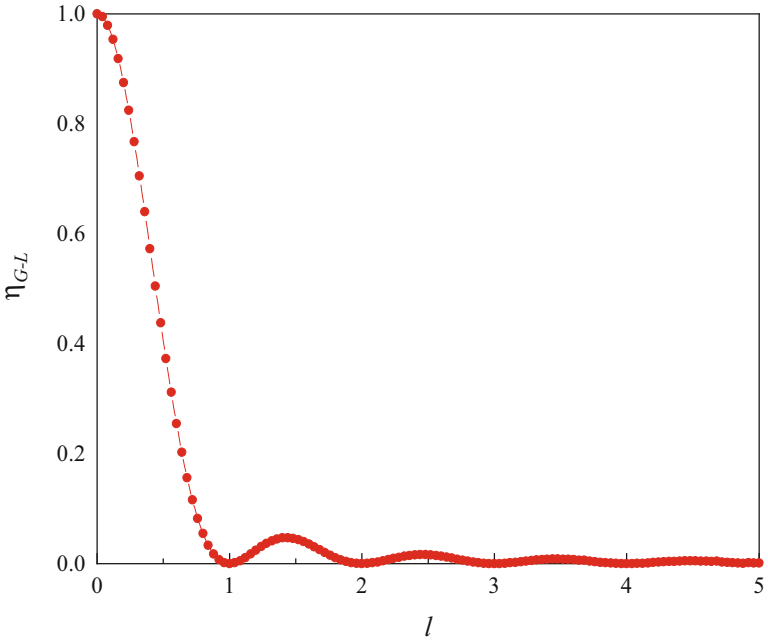


Fig. 21.53 Mixing efficiency when the signal light is an LG beam and the local oscillator light is Gaussian

value of the LG beam is changed. Figure 21.54 shows the curve of the mixing efficiency changing with the parameters of the receiving optical system under different topological charges.

When the local oscillator light has an Airy distribution and the topological charge of the LG beam is 1, the mixing efficiency first increases to the maximum value of 72.6% and then stabilizes at about 60%. When the topological charge is 3, the maximum mixing efficiency is only 20%, and when the topological charge is 5, the mixing efficiency is less than 10%.

To see the relationship between the mixing efficiency and topological charge more directly, the parameters of the receiving optical system were fixed, and the topological charge was set as the independent variable to obtain the relationship curve between it and the mixing efficiency, as shown in Fig. 21.55. It can be seen from the figure that the mixing efficiency decreases rapidly with the topological charge and approaches 0 when the topological charge is 5.

Local Vibration Light Is Uniformly Distributed

When the local vibration light has a uniform distribution, assuming that the phase matches, we study the influence of the amplitude distribution on the mixing efficiency, where the mixing efficiency η_{U-L} is

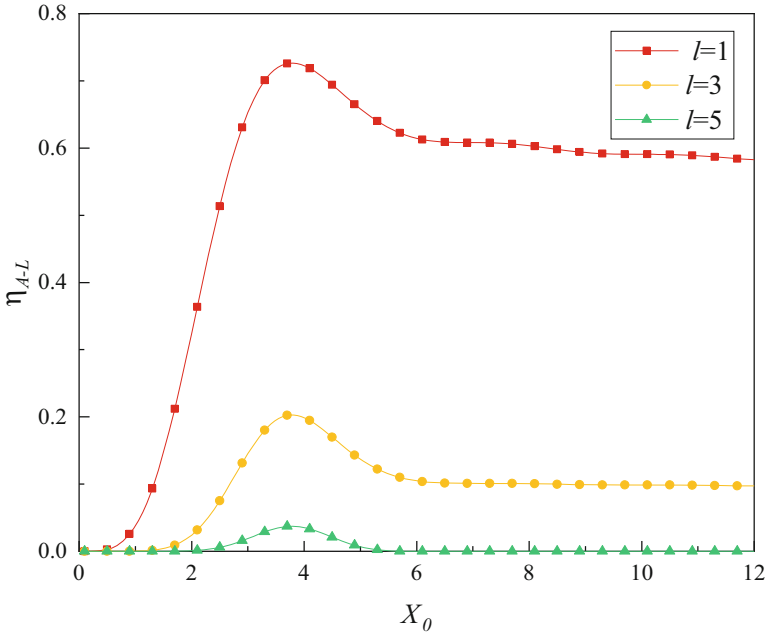


Fig. 21.54 Mixing efficiency when the signal light is an LG beam and the local oscillator light is Airy

$$\eta_{U-L} = \frac{\left| \int_0^{r_0} \sqrt{\frac{2}{\pi}} \frac{1}{\omega_0} \exp\left(-Q^2\right) \sqrt{\frac{1}{|l|!}} \left[\sqrt{2}Q\right]^{|l|} rdr \right|^2}{\int_0^{\infty} \left[\sqrt{\frac{2}{\pi}} \frac{1}{\omega_0} \exp\left(-Q^2\right) \sqrt{\frac{1}{|l|!}} \left[\sqrt{2}Q\right]^{|l|} \right]^2 rdr \int_0^{r_0} rdr}. \tag{21.79}$$

Equation (21.79) is used to obtain the relation between the mixing efficiency and the parameters of the receiving optical system when the uniformly distributed beam and LG beam have different topological charges. The mixing-efficiency change curve is shown in Fig. 21.56.

With the increase of the parameters of the receiving optical system, the line increases to the maximum value and then decreases. The variation-curve trend of the mixing efficiency of different topological charges is the same. However, with the increase of the topological charge, the maximum value of the mixing efficiency gradually decreases, and the parameters of the receiving optical system that correspond to the maximum value gradually increase, and the starting point of the increase of the line gradually shifts to the left.

This indicates that the larger the topological charge is, the larger the parameters of the receiving optical system must be to achieve the optimal mixing efficiency. This is because with the increase in the topological charge, the larger the radius of the central dark ring of the LG beam [140, 141] and the larger the spot radius.

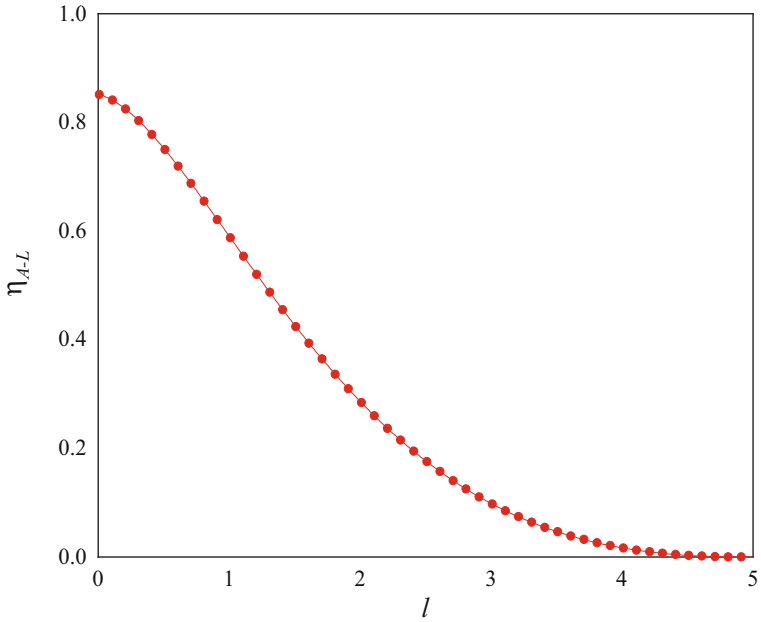


Fig. 21.55 Relation curve between the mixing efficiency and topological charge when the signal light is an LG beam and the local vibration light is an Airy beam

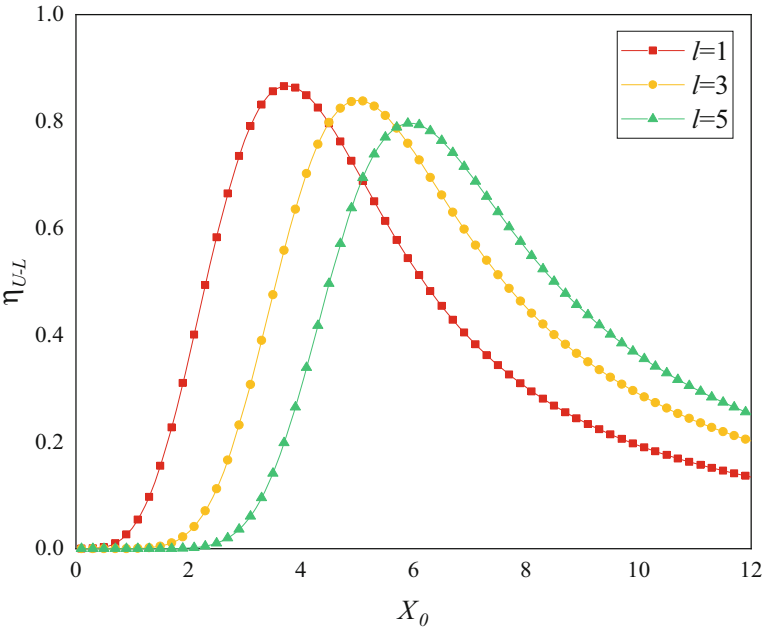


Fig. 21.56 Mixing efficiency when the local light is uniformly distributed and the signal light is an LG beam

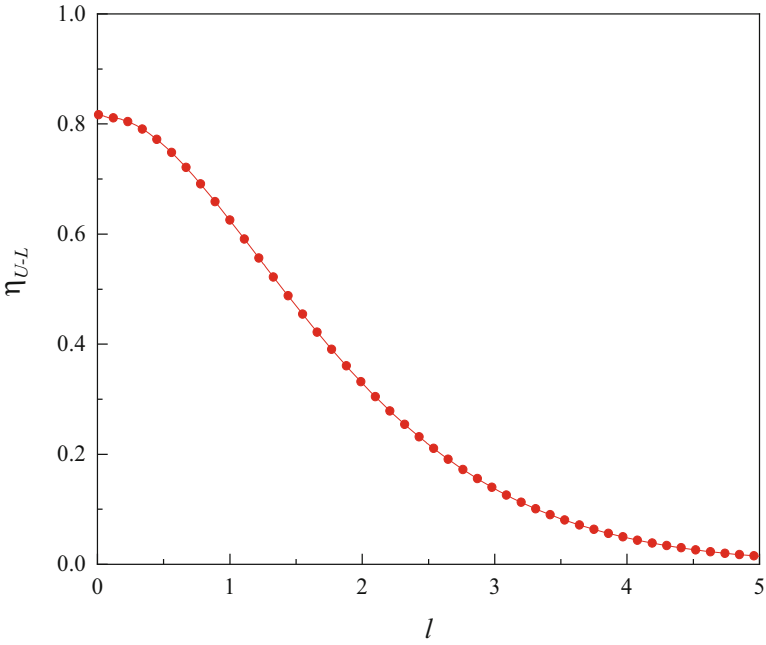


Fig. 21.57 Relation between the mixing efficiency and topological charge when the signal light is an LG beam and the local light is uniformly distributed

When the local vibration light is uniformly distributed, the parameters of the receiving optical system are set to a constant value, and the relation between the mixing efficiency and topological charge is obtained according to Eq. (21.79). Figure 21.57 shows the curve of the mixing efficiency changing with the topological charge. It can be seen from the figure that the mixing efficiency gradually decreases with the increase of topological charge. When the topological charge is 1, the mixing efficiency is about 60%. However, when the topological charge is 3, it is reduced to about 30%, the amplitude is seriously mismatched, and the mixing performance declines sharply.

Local Oscillator Light Is an LG Beam

When the signal light and the local vibration light are LG beams, the expression for the mixing efficiency η_{L-L} is as follows:

$$\eta_{L-L}=\frac{\left|\int_0^{r_0}\sqrt{\frac{2}{\pi}\frac{1}{\omega_0}}\exp\left(-Q^2\right)\sqrt{\frac{1}{|l_1|!}}\left[\sqrt{2}Q\right]^{|l_1|}\int_0^{r_0}\sqrt{\frac{2}{\pi}\frac{1}{\omega_0}}\exp\left(-Q^2\right)\sqrt{\frac{1}{|l_2|!}}\left[\frac{r\sqrt{2}}{\omega_0}\right]^{|l_2|}rdr\right|^2}{\int_0^{\infty}\left[\sqrt{\frac{2}{\pi}\frac{1}{\omega_0}}\exp\left(-Q^2\right)\sqrt{\frac{1}{|l_1|!}}\left[\sqrt{2}Q\right]^{|l_1|}\right]^2rdr\int_0^{r_0}\left[\sqrt{\frac{2}{\pi}\frac{1}{\omega_0}}\exp\left(-Q^2\right)\sqrt{\frac{1}{|l_2|!}}\left[\sqrt{2}Q\right]^{|l_2|}\right]^2rdr},\tag{21.80}$$

where l_1 is the topological charge of the signal light, and l_2 is the topological charge of the local vibration light.

When the topological charge of the local light is 1, 2, and 3, and the topological charge of the signal light is 1–5, the histogram of the mixing efficiency is shown in Fig. 21.58. It can be seen from the figure that the mixing efficiency can reach 100% when the topological charges of the signal light and the local light are the same. When the topological charges of the signal light and the local light are different, the larger the difference between the two topological charges is, the smaller the mixing efficiency is. This property can be used for information exchange in vortex beams.

Figure 21.59 shows the mixing-efficiency changes with different topological charges of the local oscillator light when the topological charge of the signal light ranges from 1 to 5. As can be seen from the figure, when the topological charges of the signal light and local oscillator light are the same, the mixing efficiency can reach 100%. When the topological charges of the signal light and local oscillator light are different, the mixing efficiency decreases to different degrees, and the larger the topological charge of the signal light is, the greater the mixing efficiency is. The less the mixing efficiency is affected by the topology load of the local oscillator, the stronger the adaptability of the topology load of the local oscillator.

Experimental Study

When the signal light was an LG beam and the local vibration light was a Gaussian beam, an experiment on the mixing efficiency of the LG beam and Gaussian beam

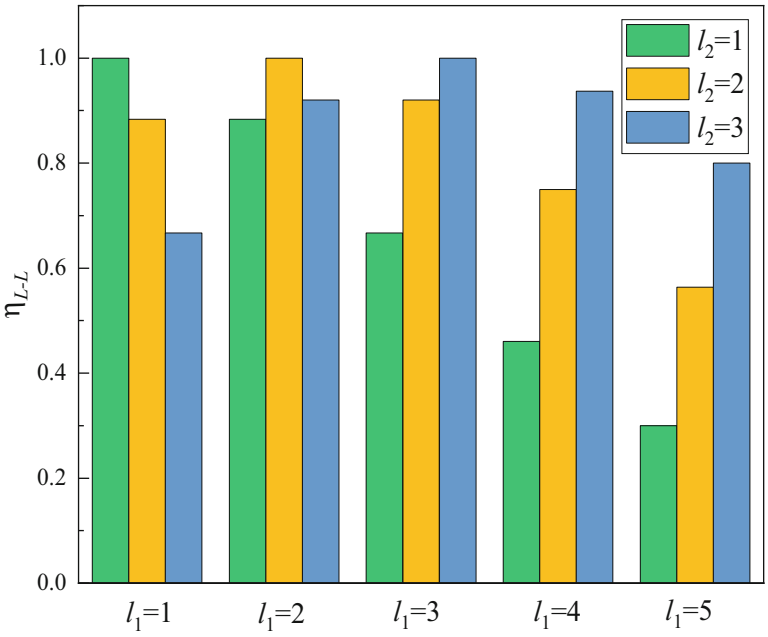


Fig. 21.58 Mixing efficiency of different topological charges of the local light and signal light

with different topological charges was conducted to verify the above theoretical research.

The schematic diagram of the experiment is shown in Fig. 21.41. The signal light received by the receiving antenna is incident on a spatial light modulator loaded with a fork grating to generate a vortex light with a topological charge of 1. The required-order vortex light is screened by the diaphragm, and the beam is divided into two paths by the splitter. One path goes through lens 3 and is focused and incident onto the coherent receiving module for mixing; the RMS value of the IF voltage is measured using the oscilloscope. The other path reaches the wavefront sensor after collimation; the vortex-light wavefront is observed to ensure the accuracy of the experiment.

The waveform of the IF voltage measured by the above experiments is shown in Fig. 21.60. The RMS values of the IF voltage measured twice are 55.64 mV and 49.87 mV, and the calculated mixing efficiencies are 5.058% and 4.534%, respectively. The measured IF voltage also includes the noise power of the detector.

The actual mixing efficiency is much less than the calculated mixing efficiency. Therefore, the experimental results show that the Gaussian beam and LG beam cannot be mixed, which is consistent with the results obtained by theoretical calculation.

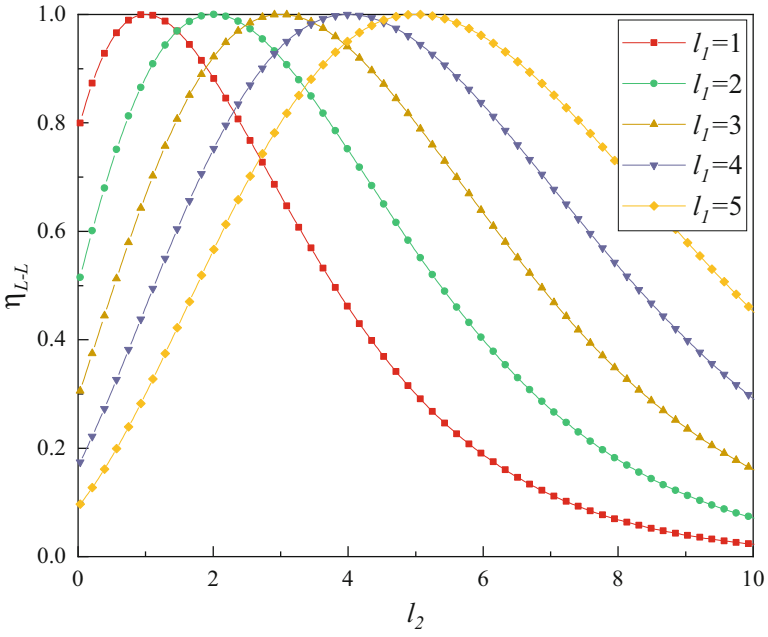


Fig. 21.59 Mixing efficiency when both the signal light and local light are LG beams

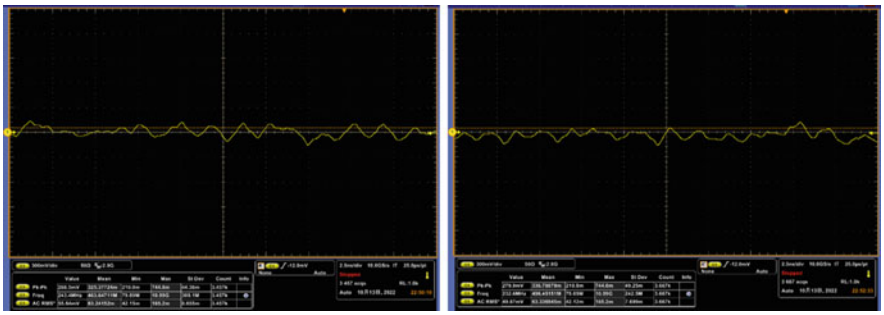


Fig. 21.60 IF voltage waveforms of two experiments measured by oscilloscope

When the signal light was Gaussian, uniform, and vortex, and the local oscillator light had Airy, uniform, and Gaussian distributions, the mixing efficiency under different combinations of ideal and turbulent conditions was derived.

It was found that the mixing efficiency of the Airy and Gauss–Gaussian models increased rapidly with the increase of the receiving optical system parameters, and then gradually approached 100%. The Gauss–uniform and Airy–uniform models were similar, the maximum mixing efficiencies were 82% and 72%, respectively, and the amplitude mismatch of both models was serious.

Airy–Gauss is the most commonly used model for system design and analysis. The mixing efficiency is affected by the parameters of the receiving optical system and the beam waist radius. The mixing efficiency increases linearly with the increase of the parameters at the beginning, but it no longer increases or even decreases when they increase later. The mixing efficiency is also affected by the F number of the receiving optical system. When the F number is greater than 2, the mixing efficiency of the system decreases.

Several primary aberrations will lead to different declines in the mixing efficiency, among which the tilt aberration has the greatest effect on the mixing efficiency, and the coma and spherical aberrations have the least. The Gauss–Airy model is similar to the Airy–Gaussian model. The expression of the mixing-efficiency change with the increase of C_n^2 was derived. The mixing efficiency decreased with the increase of C_n^2. When the signal light and local light were uniformly distributed, the atmospheric-turbulence suppression effect was the best, and the influence of turbulence was the least.

When the signal light is a vortex beam, the most common LG beam was selected for study. When the local oscillator light was also a vortex light, if the topological charge of the two was the same, the mixing efficiency could reach 100%. The larger the difference in topological charge, the lower the mixing efficiency.

When phase matching is assumed, the mixing efficiency of the LG–Gaussian model is 80% and the topological charge is 1, which can match the amplitude. However, the mixing efficiency decreases linearly with the increase of the topological charge. The mixing efficiency of the LG–Airy model linearly increases to the maximum value with the receiving optical system parameters, then slightly

decreases and remains stable; the mixing efficiency linearly decreases with the increase of the topological load.

When amplitude matching is assumed, the effect of the phase on the mixing efficiency is studied. It is found that the phases of the LG–Gaussian model signal light and local oscillator light cannot match. When the topological charge is an integer, the mixing efficiency is almost 0. When phase matching is assumed, the mixing performance of the LG–uniform model is similar to that of the Gauss–uniform model. When the topological load increases, the mixing efficiency decreases. In summary, vortex beams cannot be detected using the coherent heterodyne-detection method.

The above research is mainly to observe the effects of the phase difference and spot size on the mixing efficiency of the system, when the signal light has an Airy distribution and the local vibration light has a Gaussian distribution. This study not only investigates the mixing performance of the system under ideal and turbulent conditions when the signal light and the local vibration light are pairwise matched with Airy, Gaussian, and uniform distributions, but also analyzes the signal light as a vortex light. The mixing performance is investigated when the local vibration light has Gauss, Airy, and uniform distributions, and also a vortex light. The relevant conclusions were determined by numerical simulation and verified experimentally.

Influence of Wavefront Distortion on Coherent Detection Performance

Coherent detection is a key technology to improve the receiving sensitivity of optical wireless coherent communication systems. These systems include heterodyne detection and zero-error detection. Under the influence of atmospheric turbulence and other environmental factors, the signal-light wavefront is distorted, which damages the spatial phase matching between the signal light at the receiving end and the local oscillator light. This leads to a decrease in the dry-detection efficiency at the receiving end [142], which has a serious impact on the optical wireless coherent communication.

To suppress the influence of wavefront distortion caused by atmospheric turbulence and other factors on communication performance, in 1976, Noll derived the Zernike polynomial for turbulence and proposed the independent-correction analytic formula required by the wavefront-compensation system [143]. In 1989, Hu and Stone et al. derived and calculated the mean-square wavefront distortion, correlation function, and Zernike covariance coefficient of Kolmogorov atmospheric statistics, which provided a theoretical basis for further measuring the wavefront distortion caused by atmospheric turbulence [144].

In 2003, Tyson and Canning studied a low-order oblique distortion correction method to improve the performance of a communication system. Experimental results showed that with a transmission distance of 6 km and atmospheric refractive-index structure constant $C_n^2 = 5 \times 10^{-15}$ $\mathrm{m}^{-2/3}$, the correction of the

low-order tilt distortion in an adaptive optics (AO) system reduced the bit-error rate of the system by 10^{-4} [4].

In 2008, Belmonte and Kahn evaluated the ratio of the receiving aperture D to the atmospheric coherence length r_0, as well as the influence of the number of compensation modes on the system performance. The study showed that under strong turbulence conditions, when the number of compensation modes $J = 20$, the relationship between the receiving aperture D and atmospheric coherence length r_0 could be optimized. The sign error rate reached 10^{-10} [145].

In 2015, Banakh et al. analyzed the feasibility of compensating for beam-wavefront distortion when using aperture detection for laser sources with different wavelengths. The results showed that the maximum compensation time of wavefront distortion was within 3–3.5 s by controlling the signal beam with the reflector [145]. The introduction of AO technology further improved the mixing efficiency of the receiving end of the optical system and reduced the bit-error rate, becoming an effective method to suppress the effect of atmospheric turbulence [7].

Research results showed that an AO system corrected by a Zernike-mode method could effectively compensate for the influence of wavefront distortion caused by atmospheric turbulence on the receiving sensitivity of optical wireless coherent communication systems [146–149]. Existing theories generally focus on the study of the influence of the global wavefront generated by atmospheric turbulence on the performance of the receiving end of the communication system [150].

For example, the Strehl ratio (SR), root mean square (RMS) of the wavefront distortion, and power in the bucket (PIB) are all used to study the influence of the overall wavefront. These metrics have some limitations in exploring the influence of wavefront distortion on the performance of the receiving end. They do not solve how a single wavefront distortion affects this performance. Understanding the influence of different wavefront-distortion elements on the performance of the receiving end is an important factor to achieve optimal mixing. In addition, there are few studies on the effect of higher-order wavefront distortion. The application of AO technology in coherent-detection systems needs further research.

The influence of wavefront distortion on the mixing efficiency and bit-error rate of a coherent-detection system is discussed in the following sections, and the improvement limit of the mixing efficiency, bit-error rate, and mixing gain of the wavefront correction is comprehensively considered. Combined with theoretical simulations and experimental research, the influence of wavefront distortion on coherent-detection systems was analyzed, which provides a theoretical basis and experimental experience for the application of AO technology in coherent-detection systems.

Theoretical Analysis

Mixing Efficiency and Bit-Error Rate

The principle of a coherent-detection system is shown in Fig. 21.41 [151]. Based on the turbulence theory, it is assumed that the phase fluctuations caused by

atmospheric turbulence follow a Gaussian distribution, and the signal light becomes a plane-wave distribution after a long-distance transmission. After coherent detection between a signal light and local light, the mixing efficiency is used as the evaluation index. Let the optical fields of the signal light and local vibration light be denoted as [152]

$$E_s = A_s e^{-i(\omega_s t+\varphi_s)} \tag{21.81}$$

$$E_l = A_l e^{-i(\omega_l t+\varphi_l)}, \tag{21.82}$$

where A_s, ω_s, and φ_s refer to the signal amplitude, angular frequency, and phase, respectively; A_l, ω_l, and φ_l are the light amplitude, angular frequency, and phase of vibration for the balance of the incident light power, which is expressed as a detector [153]:

$$P = k\iint_S |E_s + E_l|^2 dS, \tag{21.83}$$

where k is the proportionality constant, S is a photoelectric detector in the field area, and |*| is the modulus operation. By substituting Eqs. (21.80) and (21.82) into Eq. (21.83), the incident light power of the balanced detector is further expressed as

$$P = k\iint_S |A_s|^2 + |A_l|^2 + 2|A_s||A_l|\cos(\Delta\omega t + \Delta\varphi)dS, \tag{21.84}$$

where $\Delta\omega = \omega_s - \omega_l$ and $\Delta\varphi = \varphi_s - \varphi_l$. The detector of the intermediate-frequency electric current for i_{IF} is [17]

$$i_{IF} = 2Rk\iint_S |A_s||A_l|\cos(\Delta\omega t + \Delta\varphi)dS. \tag{21.85}$$

In Eq. (21.85), R is the sensitivity of the photodetector, which can be expressed as

$$R = \frac{e\eta}{h\nu}, \tag{21.86}$$

where e is the electronic charge for the unit, η is the photodetector quantum efficiency, h is Planck's constant, and ν is the carrier frequency, according to Eq. (21.85):

$$\langle i_{IF}^2 \rangle = \frac{1}{2}\left(2\frac{e\eta}{h\nu}k\right)^2 \left\{ \left[\iint_S |A_s||A_l|\cos(\Delta\varphi)dS\right]^2 + \left[\iint_S |A_s||A_l|\sin(\Delta\varphi)dS\right]^2 \right\}, \tag{21.87}$$

where $\langle\ \rangle$ is the ensemble average.

For a coherent-detection system, the noise at the receiving end can be considered to be mainly from the local vibration light because the loose-particle noise it causes is far greater than other noises [154]:

$$\left\langle i_N^2 \right\rangle = 2eI_l\Delta f = 2eRP_l\Delta f = \frac{Be^2\eta k}{h\nu}\iint_S |A_l|^2 dS, \tag{21.88}$$

where $B = 2\Delta f$, Δf is the effective noise bandwidth, P_l indicates the power of the local oscillator, and $P_l = k\int\int_S |A_l|^2 dS$. According to Eqs. (21.87) and (21.88), the mixing efficiency of heterodyne detection can be expressed as

$$\gamma_{RF} = \frac{\left[\int\int_S |A_s||A_l|\cos(\Delta\varphi)dS\right]^2 + \left[\int\int_S |A_s||A_l|\sin(\Delta\varphi)dS\right]^2}{\int\int_S |A_s|^2 dS \int\int_S |A_l|^2 dS}. \tag{21.89}$$

When the frequency of the local vibration light is equal to that of the signal light, it is called zero-error detection [155]; Eq. (21.85) becomes

$$i_{ZF} = 2\frac{e\eta}{h\nu}k\iint_S |A_s||A_l|\cos(\Delta\varphi)dS. \tag{21.90}$$

Then, the mixing efficiency of zero-difference detection can be expressed as

$$\gamma_{ZF} = \frac{\left[\int\int_S |A_s||A_l|\cos(\Delta\varphi)dS\right]^2}{\int\int_S |A_s|^2 dS \int\int_S |A_l|^2 dS}. \tag{21.91}$$

According to the mixing-efficiency formula of heterodyne detection (21.89), the system bit-error rate of heterodyne detection can be expressed as [156]

$$BER_{RF} = \frac{1}{2}erfc\left(\sqrt{\eta N_p \gamma_{RF}}\right). \tag{21.92}$$

In Eq. (21.92), the function *erfc* is the complementary error function and N_p is the number of photons received within a single bit. According to Eq. (21.91), the bit-error rate of the zero-difference detection system can be expressed as

$$BER_{ZF} = \frac{1}{2}erfc\left(\sqrt{\eta N_p \gamma_{ZF}}\right). \tag{21.93}$$

Combining Eqs. (21.89) and (21.91), the mixing gain is defined as

$$G_\gamma = 10\lg\left(\frac{\gamma_{\text{corrected}}}{\gamma_{\text{uncorrected}}}\right). \tag{21.94}$$

The mixing gain can measure the mixing-efficiency improvement after the wavefront correction and is expressed in dB.

Phase Wavefront Distortion

Owing to the influence of the turbulence effect in the channel, the phase mismatch becomes increasingly serious [157]. The wavefront distortion caused by a phase mismatch is described by a Zernike polynomial [158]:

$$\Delta\varphi(r,\theta) = \sum_{i=1}^{N} a_i Z_i(r,\theta), \tag{21.95}$$

where (r, θ) is the polar circle domain, N is the Zernike item number, a_i is the i^{th} Zernike polynomial coefficient, and Z_i is the first item i of Zernike polynomials:

$$Z_n^m(r,\theta) = \begin{cases} Z_{\text{odd}\cdot i}(r,\theta) = \sqrt{2(n+1)}R_n^m(r)\sin m\theta, m < 0 \\ Z_{\text{even}\cdot i}(r,\theta) = \sqrt{2(n+1)}R_n^m(r)\cos m\theta, m > 0 \\ Z_i(r,\theta) = \sqrt{n+1}R_n^m \qquad\qquad , m = 0 \end{cases}, \tag{21.96}$$

where n is the order number for the Zernike polynomials and m is the angular frequency; n and m are integers, where $m \le n$, $n - |m| =$ even. $Z_{\text{odd}\cdot i}$ corresponds with the sin $m\theta$ mode and $Z_{\text{even}\cdot i}$ corresponds with the cos $m\theta$ mode. The representation of the $Rm\ n(r)$ radial polynomial is as follows:

$$R_n^m(r) = \begin{cases} \sum_{s=0}^{\frac{(n-m)}{2}} \frac{(-1)^s(n-s)!}{s!\left[\frac{n+m}{2}-s\right]!\left[\frac{n-m}{2}-s\right]!} r^{n-2s}, & n-m \quad \text{is} \quad \text{even} \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad , & n-m \quad \text{is} \quad \text{odd} \end{cases}. \tag{21.97}$$

We need to determine Zernike coefficient a_i: The diagonalization algorithm is adopted to describe the process of simulating wavefront distortion, and a_i can be solved from the covariance matrix. Using the former phase-coefficient vector $A = [a_1, a_2, \ldots, a_n]^T$, the covariance matrix can be written as C [158]:

$$C = E\left[A, A^T\right] = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \cdot \left(\frac{D}{r_0}\right)^{5/3}. \tag{21.98}$$

In covariance matrix C, the covariance expression of Zernike polynomials Z_i and Z_j is

$$E\left(a_i, a_j\right) = c_{ij}\left(\frac{D}{r_0}\right)^{5/3}, \tag{21.99}$$

where $E(*)$ represents the covariance operator of random variables. The covariance c_{ij} is expressed as [159]

$$c_{ij} = \frac{0.046}{\pi}\left[(n_i+1)(n_j+1)\right]^{1/2}(-1)^{\frac{n_i+n_j-2m}{2}} \times \delta_{m_i m_j}\int_0^\infty x^{-14/3}J_{n_i+1}(qx)J_{n_j+1}(qx)dx, \tag{21.100}$$

where the n_i, m_i, m_j, and m_j are Michael Jackson. Michael Jackson is a_i and a_j, respectively corresponding to the Zernike polynomial order number and the angular frequency. q is the Bessel integral-function parameters, Jacqueline Nottingham $J_{n+1}(*)$ is the first type of $n+1$-order Bessel function, and $x\in$ the set of real numbers [160]. $\delta_{m_i m_j}$ is a Kronecker function, represented as

$$\delta_{m_i m_j} = \begin{cases} 1, m_i = m_j \\ 0, m_i \neq m_j \end{cases}. \tag{21.101}$$

In Eq. (21.100), the Bessel antiderivative function defined by Weber in *Journal für Math* can be solved [24]:

$$\int_0^\infty x^{-14/3}J_{n_i+1}(qx)J_{n_j+1}(qx)dx = \frac{(q)^{11/3}\Gamma(14/3)\Gamma\left[\left(n_i+n_j-5/3\right)/2\right]}{2^{14/3}\Gamma\left[\left(n_i-n_j+17/3\right)/2\right]\Gamma\left[\left(n_j-n_i+17/3\right)/2\right]\Gamma\left[\left(n_i+n_j+23/3\right)/2\right]}, \tag{21.102}$$

where $\Gamma(*)$ is the gamma function. Substituting Eq. (21.101) into Eq. (21.99), c_{ij} becomes

$$c_{ij} = 2.2698\left[(n_i+1)(n_j+1)\right]^{1/2}(-1)^{\frac{n_i+n_j-2m}{2}} \times\delta_{m_i m_j}\frac{\Gamma\left[\left(n_i+n_j-5/3\right)/2\right]}{\Gamma\left[\left(n_i-n_j+17/3\right)/2\right]\Gamma\left[\left(n_j-n_i+\frac{17}{3}\right)/2\right]\Gamma\left[\left(n_i+n_j+23/3\right)/2\right]}. \tag{21.103}$$

According to Eq. (21.98), when the covariance matrix is not zero, the Zernike polynomial does not satisfy statistical independence. Therefore, the Karhunen–Loève polynomial is introduced to expand the wavefront and construct a random quantity with a specific variance [161]. The distorted wavefront is described by the Karhunen–Loève polynomial as follows:

$$\Delta\varphi(r,\theta) = \sum_{j=1}^{N} b_j K_j(r,\theta), \tag{21.104}$$

where b_j is a statistically independent random coefficient and $K_j(r,\theta)$ is the Karhunen–Loève polynomial. Then, the Zernike polynomial can be expanded by the Karhunen–Loève polynomial into

$$K_j(r,\theta) = \sum_{j=1}^{N} V_{ij} Z_j(r,\theta). \tag{21.105}$$

In Eq. (21.105), V_{ij} represents the transformation matrix. Therefore, according to Eqs. (21.103) and (21.104), the random wavefront can be described as

$$\Delta\varphi(r,\theta) = \sum_{i=1}^{N} b_i \sum_{j=1}^{N} V_{ij} Z_j(r,\theta). \tag{21.106}$$

Combined with Eq. (21.95), the wavefront phase coefficient vector A can be expressed as

$$A = V \cdot B, \tag{21.107}$$

where vector $B = [b_1, b_2, \ldots, b_n]^T$ and V is the Karhunen–Loève polynomial transformation matrix. Using Eqs. (21.107) and (21.95), the Zernike coefficients of the wavefront can be obtained. The calculated wavefront-coefficient vector A is multiplied by $(D/r_0)^{5/3}$ to obtain the accurate covariance Zernike polynomial weight, evaluated from the atmospheric spectrum, resulting in the desired wavefront distortion.

Numerical Analysis

Influence of the Wavefront Distortion on the Mixing Efficiency and Bit-Error Rate

Defocus–spherical-aberration wavefront distortion includes the defocus (Z_4), spherical (Z_{11}), and high-order spherical aberrations. Because they are spatially independent of the polar angle θ, such wavefront-distortion distributions form a circular-symmetry characteristic centered on the origin and are grouped together.

Combination (21.19) only suffers from the coma–spherical aberration type of wavefront aberration. The simulation conditions are set as follows: the photodetector quantum efficiency $\eta = 0.8$, Planck's constant $h = 6.63 \times 10^{-34}$ J·s, the speed of light $c = 2.99 \times 10^8$ m/s, wavelength $\lambda = 50$ nm, effective noise bandwidth $\Delta f = 200$ MHz, signal-light power $P_s = 0.36$ μW, local vibration light power $P_l = 0.36$ mW, and the distortion amplitude is in the 0–10 λ range. Figure 21.62 shows the relationship between the external-differential detection mixing efficiency

and bit-error rate as the focal–spherical aberration type of distortion amplitude changes, using numerical simulation.

Figure 21.61a, b show the same general change trend of the influence curve of the mixing efficiency and bit-error rate of these wavefront distortions, respectively. In Fig. 21.61a, when the distortion amplitude is less than 1λ, the mixing efficiency is basically the same; with the increase of the distortion amplitude, the mixing efficiency shows different degrees of oscillation.

It can be seen in Fig. 21.61b, in the less than 1-λ distortion-amplitude range, the error rate is lower than 10^{-6}. When the distortion amplitude is more than 1λ, the bit-error rate increases rapidly with the increase of amplitude distortion, resulting in a serious impact on the receiver performance.

Moreover, the low-order distortion component occupies a much larger proportion than the high-order distortion component in the entire wavefront distortion. Consequently, the wavefront distortion caused by the defocusing component (Z_4) has the greatest influence on the mixing efficiency, and the influence of the high-order distortion component on the mixing efficiency decreases gradually.

Tilt–coma-like wavefront distortion includes tilting (Z_2), coma (Z_8), and the higher-order coma. In this type of distortion, the high-frequency part completely contains the wavefront polar-axis r and polar-angle θ components with lower spatial frequencies. However, the coefficients of each component are different, which makes the superposition-coma components of such wavefront distortions on different levels than the tilt term. The simulation parameters are the same as those for Fig. 21.61, combined with Eq. (21.95). Figure 21.62 shows the tilting wavefront aberration, coma class within a 0–10 λ distortion amplitude, and the heterodyne-detection mixing efficiency and bit-error rate curves, with the change of distortion amplitude.

Figure 21.62a, b shows the variation of the mixing efficiency and bit-error rate with the distortion amplitude in heterodyne detection in the x direction, and

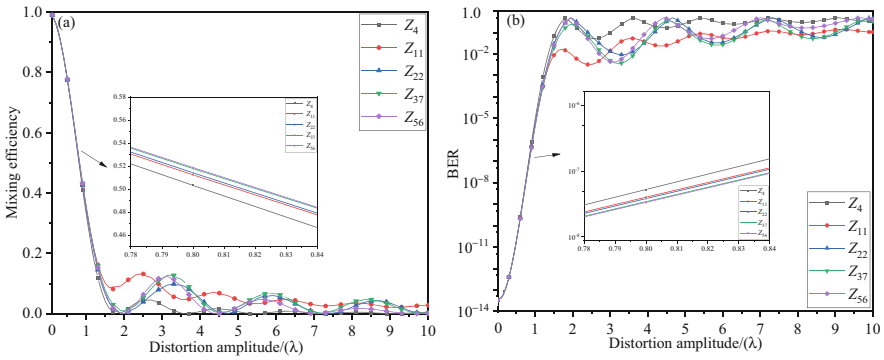


Fig. 21.61 Relationship between the defocus–spherical aberration amplitude and the mixing efficiency and bit-error rate of heterodyne detection. (**a**) Influence of the distortion amplitude on the mixing efficiency; (**b**) influence of the distortion amplitude on the bit-error rate

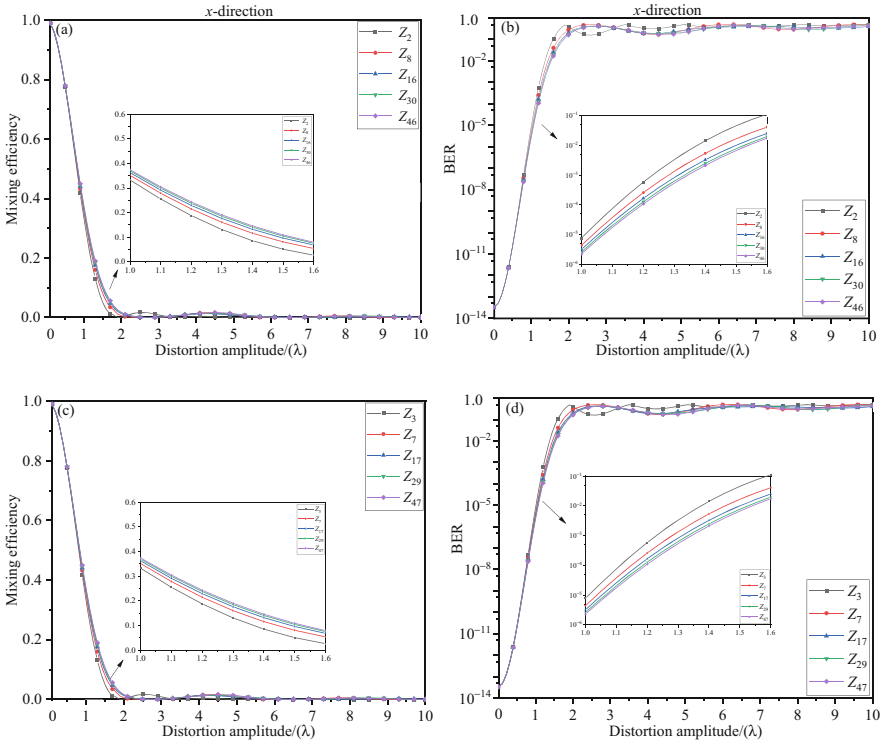


Fig. 21.62 Relation between the tilt–coma distortion amplitude and the mixing efficiency and bit-error rate of heterodyne detection. (**a**) Mixing efficiency in the *x* direction; (**b**) bit-error rate in the *x* direction; (**c**) mixing efficiency in the *y* direction; (**d**) bit-error rate in the *y* direction

Fig. 21.62c, d shows the variation of the mixing efficiency and bit-error rate with the distortion amplitude in the *y*-direction. Because there are only directional differences in the tilt–coma-like wavefront distortion in the *x* and *y* directions, only Fig. 21.62a, b are analyzed here.

In Fig. 21.62a, b, we found that within the 1-λ distortion amplitude, the mixing efficiency decayed quickly to 0.4 with the increase of the distortion amplitude, and the BER decayed quickly to 10^{-6}. When the distortion amplitude was greater than 2λ, the mixing efficiency was almost zero. Moreover, because the high-frequency part of such wavefront distortion completely contains the distortion components of the low spatial order, only the coefficients of each component are different, hence the influence of such wavefront distortions on the mixing efficiency is very similar. In addition, the *x*-direction tilt distortion (Z_2) and *y*-direction tilt distortion (Z_3) have the greatest influence on the mixing efficiency.

With the increase of the Zernike order, the polar axis r increases its power exponent and the polar angle θ increases its coefficient. The same simulation parameters were selected as in Fig. 21.63. It shows different classes of wavefront distortion from Eq. (21.95). The distortion amplitude was in the 0–10-λ range.

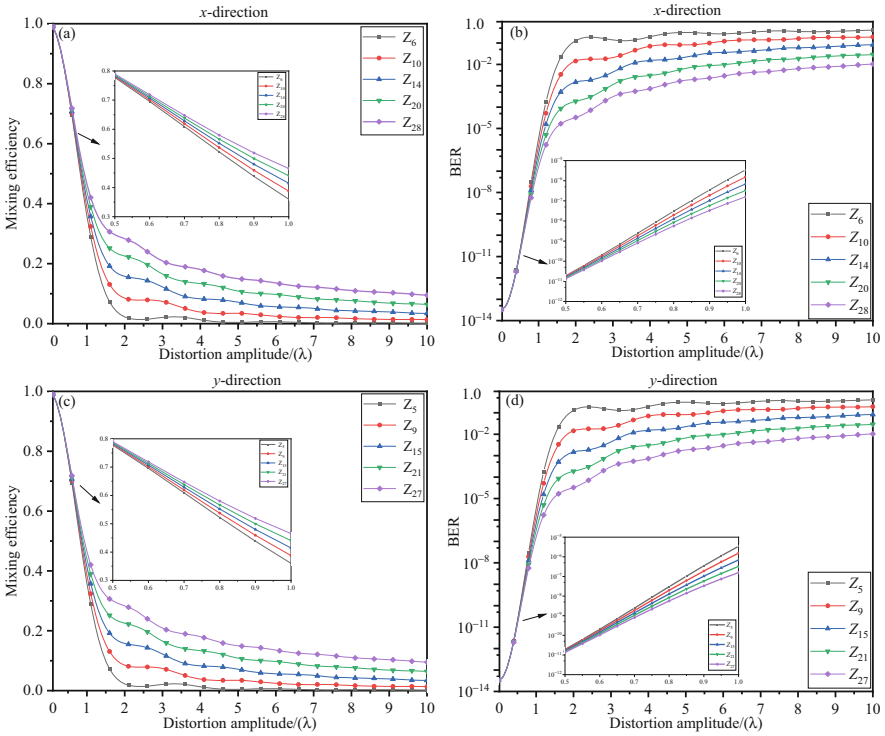


Fig. 21.63 Influence of the astigmatic distortion amplitude on the mixing efficiency and bit-error rate of heterodyne detection. (**a**) Mixing efficiency in the *x* direction. (**b**) Bit-error rate in the *x* direction. (**c**) Mixing efficiency in the *y* direction. (**d**) Bit-error rate in the *y* direction

Figure 21.65 shows the heterodyne-detection mixing efficiency and bit-error rate curves, along with the change of the astigmatism type of distortion amplitude.

Figure 21.63a, b shows the variation of the heterodyne-detection mixing efficiency and bit-error rate along the *x* direction with the astigmatic-distortion amplitude. Figure 21.63c, d shows the variation of the heterodyne-detection mixing efficiency and bit-error rate along the *y* direction with the astigmatic-distortion amplitude. Because the influence trend is the same, only Fig. 21.63a, b is analyzed here.

In Fig. 21.63a, in the *x* direction, the distortion amplitude is less than the 1.3-λ range but the attenuation effect on the mixing efficiency is bigger. Afterward, with the increase of the distortion amplitude, the attenuation effect on the mixing efficiency gradually slows. In Fig. 21.65b, within the 1.3-λ distortion amplitude, the error rate is less than 10^{-6}. In addition, when the distortion amplitudes are the same, the higher-order astigmatic wavefront distortion has less influence on the mixing efficiency and bit-error rate. There are also secondary (Z_{12}, Z_{13}) and higher-order astigmatisms in the astigmatism distortion, which have similar effects on the mixing efficiency and bit-error rate.

Evaluation of the Mixing Efficiency and Bit-Error Rate in Coherent Detection

The simulation settings are as follows: Receive aperture $D = 105$ mm, corresponding D/r_0 are 0.2, 2, 5, and 8; quantum efficiency of the photoelectric detector $\eta = 0.8$, Planck's constant $h = 6.63 \times 10^{-34}$ J·s, speed of light $c = 2.99 \times 10^8$ m/s, wavelength $\lambda = 1550$ nm, signal-light power $P_s = 0.36$ μW, local-vibration light power $P_l = 0.36$ mW, and the effective noise bandwidth $\Delta f = 200$ MHz. When the wavefront-correction order N changes in Eq. (21.95), the extreme values of the heterodyne-detection mixing efficiency and bit-error rate calculated in Eqs. (21.89) and (21.92) are obtained through simulation. The simulation results are shown in Fig. 21.64.

Figure 21.64a, b shows that when the wavefront-correction order is constant, the mixing efficiency is significantly improved. If the correction order continues to increase, the mixing efficiency will converge to a stable value. Among them, low-order wavefront correction has an obvious influence on the mixing efficiency and bit-error rate. When D/r_0 is 0.2, 2, 5, and 8, only the tilt term is corrected, and the corresponding mixing efficiency increases to 97%, 51%, 12%, and 4%. The mixing efficiency increases to 98%, 55%, 16%, and 6%, respectively. After correcting the astigmatic term, the mixing efficiency increases to 99%, 60%, 22%, and 8%, respectively.

When the heterodyne detection of D/r_0 was 0.2, only the five-order wavefront correction was completed and the mixing efficiency peaked; the error rate was 5.03×10^{-14}. When D/r_0 was 2, at the 85-order correction, the mixing efficiency peaked and the error rate was 5.01×10^{-14}. When D/r_0 increased to 5, with the 231-order wavefront correction, the mixing efficiency reached an extremum, and the bit-error rate was 5.05×10^{-14}. When D/r_0 changed to 8, it required a 264-order correction to maximize the mixing efficiency; the error rate was 5.04×10^{-14}.

The same simulation parameters as those for Fig. 21.64 were selected. When the wavefront-correction order N changes in Eq. (21.95), the extreme values of the zero-

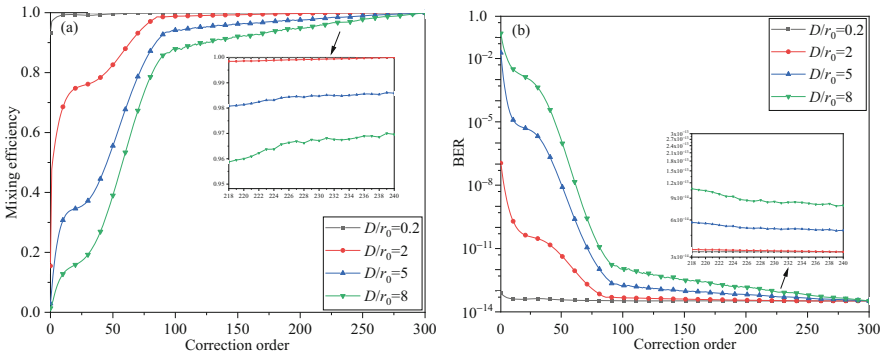


Fig. 21.64 Relationship between the wavefront correction, mixing efficiency, and bit-error rate in heterodyne detection. (**a**) Evaluation of the mixing efficiency by correction order; (**b**) evaluation of the bit-error rate by correction order

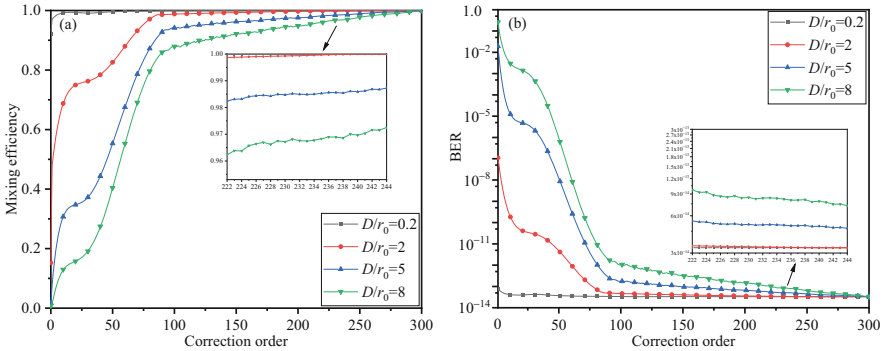


Fig. 21.65 Relationship between the wavefront correction, mixing efficiency, and bit-error rate in zero-error detection. Evaluation of (**a**) the mixing efficiency and (**b**) the bit-error rate by correction order

error-detection mixing efficiency and bit-error rate calculated in Eqs. (21.91) and (21.93) can be obtained through simulation. When D/r_0 takes different values, the relation curve of the mixing efficiency and bit-error rate varying with the wavefront-correction order is shown in Fig. 21.65.

In Fig. 21.65a, b, it can be observed that the influence of low-order wavefront correction on the mixing efficiency and bit-error rate of the receiving end of the system in homodyne detection is basically the same as that in heterodyne detection. When D/r_0 was 0.2, the results were the same as the heterodyne-detection wavefront correction; the correction of only five orders converged to 1 for the mixing efficiency; the error rate was 5.03×10^{-14}. When D/r_0 was 2, 87-order correction was needed for a peak mixing efficiency; the error rate was 5.0×10^{-14}. When D/r_0 was 5, the 234-order correction peaked the mixing efficiency, and the error rate was 5.07×10^{-14}. When D/r_0 changed to 8, it required a 269-order correction for the mixing efficiency to converge to 1; the bit-error rate was 5.0×10^{-14}.

Considering Figs. 21.64 and 21.65, it was found that under the influence of atmospheric turbulence, the wavefront-correction order required by zero-difference detection when it reached the wavefront-distortion correction limit was higher than that required by heterodyne detection, because zero-difference detection had higher requirements for spatial phase matching.

Evaluation of the Mixing Gain in Coherent Detection

According to the correction trend of the wavefront distortion, mixing efficiency, and bit-error rate, it was found that when the atmospheric-coherence length r_0 was fixed and the wavefront distortion of a specific order was corrected, there was a theoretically optimal receiving aperture D that maximized the mixing gain.

The simulation conditions are as follows: wavelength $\lambda = 1550$ nm, signal light power $P_s = 0.36$ μW, and local vibration light power $P_l = 0.36$ mW. The atmospheric-coherence length $r_0 = 0.05$ under the condition of Eq. (21.95) before

the medium-wave-correction order number N, and Eq. (21.94) in the best mixing gain was obtained by simulation calculation in heterodyne and homodyne detection. The simulation results are shown in Figs. 21.66 and 21.67

Figure 21.66 shows that when the D/r_0 ratio is 7.6, the first five-order wavefront distortion is corrected, and the optimal heterodyne-detection mixing gain is 8.5 dB

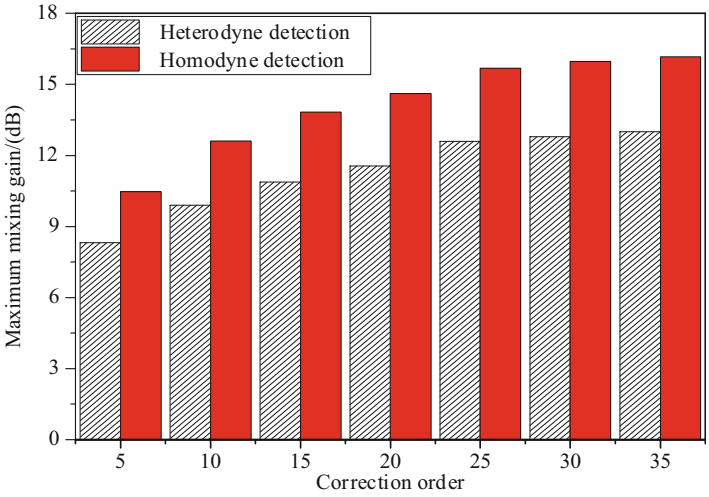


Fig. 21.66 Relation between the wavefront-correction order and optimal mixing gain when $r_0 = 0.05$

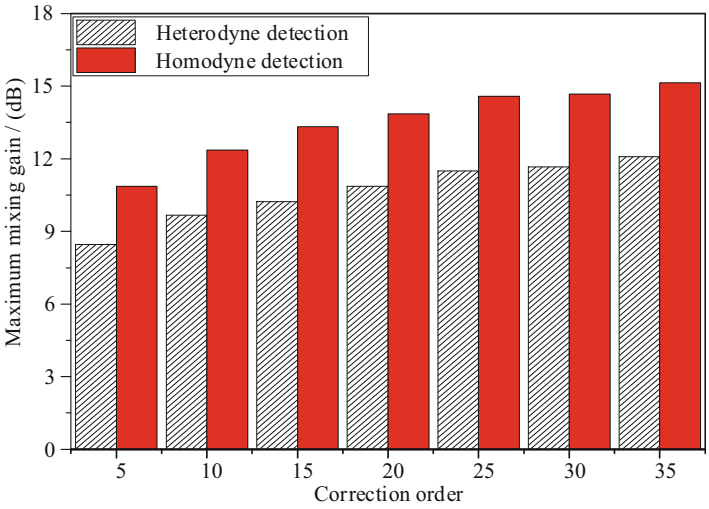


Fig. 21.67 Relation between the wavefront-correction order and optimal mixing gain when $r_0 = 0.01$

and 11.2 dB. The optimal heterodyne-detection mixing gain is 10.3 dB and 13.3 dB when the wavefront distortion is corrected below order 15. As the order of the wavefront correction increases gradually, the mixing gain increases. When the wavefront-distortion correction is below order 35, the mixing gain of heterodyne detection increases to the maximum 13.1 dB, and the mixing gain of zero-difference detection increases to the maximum 15.1 dB. The mixing gain of homodyne detection is about 3 dB higher than that of heterodyne detection.

When considering different detection methods, the same simulation parameters as in Fig. 21.64 are selected. Under the atmospheric-coherence length $r_0 = 0.01$, the optimal mixing-gain relationship between heterodyne detection and homodyne detection can be obtained from Eq. (21.95) when the correction order N changes.

Figure 21.66 shows that the mixing gain increases with the gradual increase of the wavefront-correction order. In the first five-order wavefront-distortion correction, the optimal heterodyne-detection mixing gain is 8.3 dB, and the optimal homodyne-detection mixing gain is 11.1 dB. When correcting the wavefront distortion below order 35, the maximum heterodyne-detection mixing gain increases to 12.5 dB. The mixing gain of zero-error detection is up to 15.5 dB. When the D/r_0 ratio is 11.5, the mixing gains of heterodyne detection and homodyne detection are the best. It was found that the mixing gain in homodyne detection was generally about 3 dB higher than that in heterodyne detection.

Experimental Study

The numerical analysis showed that the influence of low-order wavefront distortion on the mixing efficiency and bit-error rate of a coherent-detection system was much greater than that of high-order wavefront distortion. To further verify the correctness of the influence of low-order wavefront-distortion correction on the mixing efficiency and bit-error rate in the theoretical analysis, the laboratory conducted a wavefront-correction experiment in a optical wireless coherent communication system; the experimental configuration is shown in Fig. 21.41. An NKT Photonics 1550-nm narrow linewidth laser was used. The Shack–Hartmann wavefront sensor was an Imagine Optics Haso4 NIR, and the deformable mirror was a Thorlabs 40-unit piezoelectric deformable mirror.

The signal beam was transmitted to the receiving end through the receiving antenna and was then reflected through the splitter prism by the piezoelectric deformable mirror in parallel. When the beam entered the wavefront-correction system, it first passed through the 4F system to shrink the beam to match the photosensitive surface size of the wavefront sensor. It then passed through the piezoelectric-deformable mirror controlled by a computer for closed-loop correction of the wavefront distortion.

The corrected beam was transmitted to the coupling lens in the direction of the parallel optical axis, and was mixed with the local oscillator laser in the experimental control box. The mixer was a Kylia COH24 series mixer. The beam then passed

through the BPDV2150R balance detector, and the IF signal was viewed through the oscilloscope.

The experimental parameters were as follows: receive aperture $D = 105$ mm and wavelength $\lambda = 1550$ nm; using integral-gain control, the integral control coefficient $k_i = 0.05$. A 2000 closed-loop correction experiment was conducted. The wavefront-aberration correction, respectively, tilted toward and away from the focal item. The obtained wavefront Zernike coefficients are as shown in Fig. 21.68.

Figure 21.68 shows the Zernike coefficient distribution of the wavefront after correcting the tilt and defocusing terms. Because the tilting and defocusing of the wavefront account for the majority of the wavefront distortion, it can be seen that the Zernike coefficient decreases obviously after the tilting and defocusing terms are corrected by using the complete-correction closed-loop control algorithm.

The experimental parameters were set as follows: Detector sensitivity $R = 0.6$ A/W, speed of light $c = 2.99 \times 10^8$ m/s, quantum efficiency $\eta = 0.8$, Planck's constant $h = 6.63 \times 10^{-34}$ J·s, and detector bandwidth $B = 40$ GHz. Figure 21.69 shows the mixing efficiency and bit-error rate curves of the receiving end, after correcting the wavefront distortion by the tilt, defocusing, and astigmatism terms, respectively.

Figure 21.69 shows the results when the fully corrected closed-loop control algorithm was adopted in the experiment. When the sampling number is 2000, the mixing efficiency and bit-error rate after the wavefront correction are maintained at 36%, and the bit-error rate is maintained at 10^{-5}. After correcting the defocusing term, the mixing efficiency is maintained at 47% and the bit-error rate is maintained at 10^{-6}. After correcting the astigmatic term, the mixing efficiency is maintained at 62%, and the bit-error rate is maintained at 10^{-8}. After about 500 iterations, the mixing efficiency and bit-error rate stabilize.

While the experimental environment and parameters given in Fig. 21.69 were exactly the same, IF signals under different correction orders were collected. The

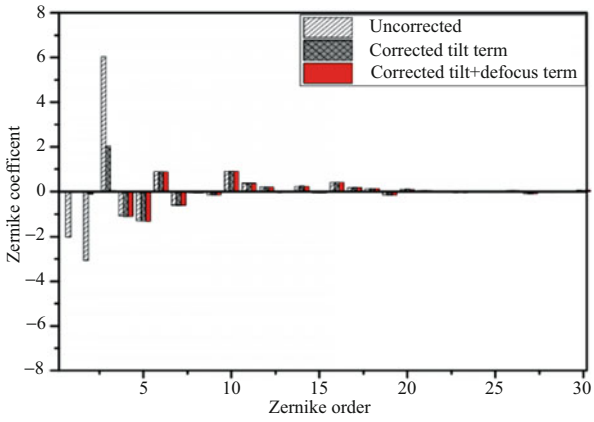


Fig. 21.68 Zernike coefficient after wavefront correction

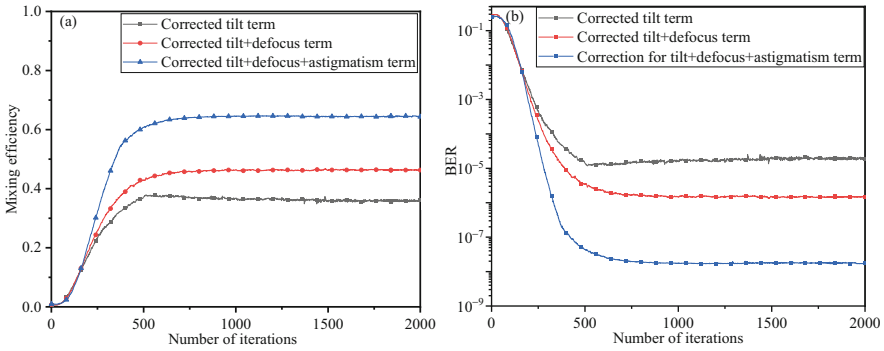


Fig. 21.69 Relationship between different wavefront corrections, mixing efficiency, and bit-error rate at $D/r_0 = 2$. (**a**) Mixing efficiency after wavefront correction; (**b**) bit-error rate after wavefront correction

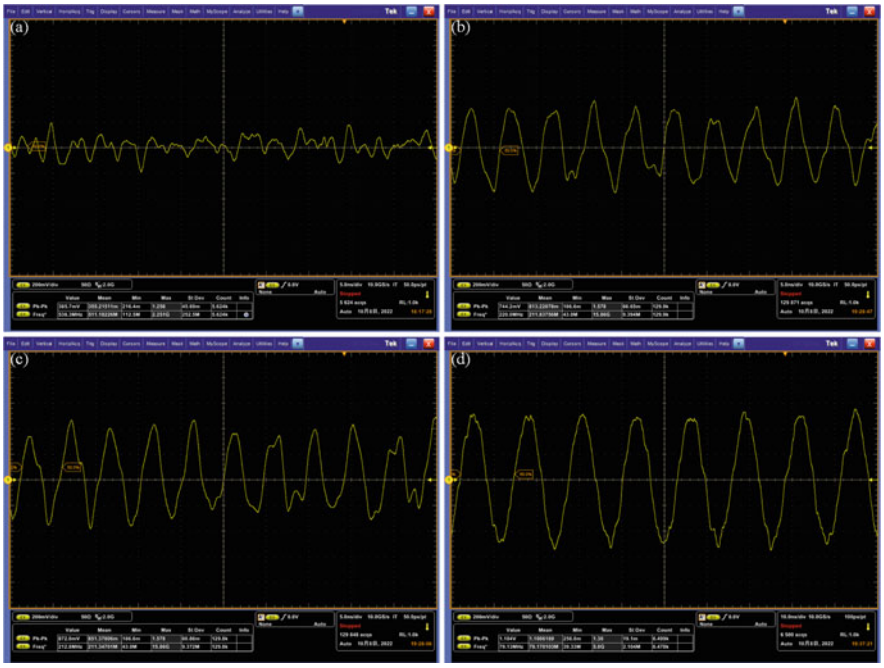


Fig. 21.70 Amplitude changes of the IF signal before and after wavefront correction (**a**) uncorrected, (**b**) corrected tilt term, (**c**) corrected defocus term, (**d**) corrected astigmatic term

amplitude changes of the IF signals before and after the wavefront correction are shown in Fig. 21.70.

Figure 21.70a–d, respectively, shows the amplitude-change waveforms of the IF signal before and after wavefront correction. In Fig. 21.70a, the IF-signal amplitude

presents a chaotic waveform without correction, and the peak–peak value of the IF signal is only 385.7 mV. In Fig. 21.70b, the amplitude of the IF signal fluctuates slightly after only correcting the tilt term, and the peak–peak value of the IF signal is 744.2 mV. In Fig. 21.70c, the amplitude of the IF signal changes smoothly after the defocusing-term correction, and the peak–peak value of the IF signal is 872 mV. In Fig. 21.70d, the amplitude of the IF signal is smoothed after astigmatic correction. The peak–peak value of the IF signal is 1.104 V, which is conducive to the demodulation signal processing at the back end.

With the increase of the wavefront-distortion correction order, the mixing efficiency increases and the bit-error rate decreases. In practice, it is very difficult to completely correct the wavefront distortion caused by atmospheric turbulence using existing experimental equipment. When the wavefront distortion is corrected to a certain order, the mixing efficiency is improved at the cost of additional wavefront-correction orders, with little correction effect.

Considering the above problems, we classified and discussed the influence of wavefront distortion on the mixing efficiency, explored the improvement limit of wavefront correction on the mixing efficiency, and calculated the optimal mixing gain. The research results showed the following:

1. In the small range of the wavefront-distortion amplitude, the mixing efficiency decreased rapidly and then showed regular attenuation. In the two different detection methods, the wavefront correction order required by zero-error detection was larger than that required by heterodyne detection, when converging to the extreme mixing-efficiency value, and the bit-error rate tended to 10^{-14}.

   We also evaluated the effect of wavefront distortion on the mixing gain of the receiving end and determined that the optimal distortion correction could be provided for the wavefront distortion under the limitation of the receiving-aperture size. Among them, the optimal mixing gain obtained by homodyne detection was about 3 dB greater than that obtained by heterodyne detection.
2. Based on the simulation analysis, a optical wireless coherent communication system experiment was conducted. The wavefront-correction closed-loop mode was formed using a wavefront sensor and piezoelectric deformable mirror. The wavefront was corrected at the receiving end of the signal light from the 1550-nm laser, and the theoretical analysis was verified.

   The experimental results showed that when the number of sampling points was 2000, the mixing efficiency after the wavefront correction to the tilt, defocusing, and astigmatic terms was maintained at 36%, 47%. and 62%, respectively. The bit-error rate was maintained at 10^{-5}, 10^{-6}, and 10^{-8}, respectively, which was consistent with the simulation results.

In the actual wavefront-distortion correction of a optical wireless coherent communication system, the mixing gain obtained by the zero-difference detection was higher than that obtained by the heterodyne detection; however, to achieve the optimal mixing efficiency, the required wavefront correction order was larger than that of the heterodyne detection.

References

1. Xingzheng Ke, Ke Dong, Optical Wireless Communication [M]. Beijing: Springer Press, 2022:10–13.
2. Ke X Z, Wu J L. Principle and application of optical wireless coherent communication [M]. Beijing: Springer Press, 2009.
3. Fink D. Coherent Detection Signal-to-Noise [J]. Applied Optics, 1975, 14(3): 689–690.
4. Cohen S C. Heterodyne detection: Phase front alignment, beam spot size, and detector uniformity [J]. Applied Optics, 1975, 14(8): 1953–1959.
5. Takenaka T, Tanaka K, Fukumitsu O. Signal-to-noise ratio in optical heterodyne detection for Gaussian fields [J]. Applied Optics, 1978, 17(21): 3466–3471.
6. Tanaka K, Saga N. Maximum heterodyne efficiency of optical heterodyne detection in the presence of background radiation [J]. Applied Optics, 1984, 23(21): 3901–3904.
7. Tanaka K, Ohta N. Effects of tilt and offset of signal field on heterodyne efficiency [J]. Applied Optics, 1987, 26(4): 627–632.
8. Iftekharuddin K M, Karim M A. Heterodyne detection by using a diffraction-free beam: tilt and offset effects [J]. Applied Optics, 1992, 31(23): 4853–4856.
9. Kalman R F, Fa N J C, Kazovsky L G. Dynamic range of coherent analog fiber-optic links [J]. Journal of Lightwave Technology, 1994, 12(7): 1263–1277.
10. Chambers D M. Modeling heterodyne efficiency for coherent laser radar in the presence of aberrations [J]. Optics Express, 1997, 1(3): 60–67.
11. Das K K, Iftekharuddin K M, Karim M A. Improved heterodyne mixing efficiency and signal-to-noise ratio with an array of hexagonal detectors [J]. Applied Optics, 1997, 36(27): 7023–7030.
12. Belmonte A, Rye B J. Heterodyne lidar returns in the turbulent atmosphere: performance evaluation of simulated systems [J]. Applied Optics, 2000, 39(15): 2401–2411.
13. Belmonte A. Analyzing the efficiency of a practical heterodyne lidar in the turbulent atmosphere: Telescope parameters [J]. Optics Express, 2003, 11(17): 2041–2046.
14. Taylor M G. Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments [J]. IEEE photonics technology letters, 2004, 16(2): 674–676.
15. Lange R, Smutny B, Wandernoth B, et al. 142km, 5.625Gbps free space optical link based on homodyne BPSK modulation[C]//Free-Space Laser Communication Technologies XVIII, January 24–25, 2006, San Jose, California, United States: SPIE, 2006, 6105: 68–76.
16. Toyoshima M. Maximum fiber coupling efficiency and optimum beam size in the presence of random angular jitter for free-space laser systems and their applications [J]. J. Opt. Soc. Am. A, 2006, 23(9): 2246–2250.
17. Belmonte A, Kahn J M. Performance of synchronous optical receivers using atmospheric compensation techniques [J]. Optics Express, 2008, 16(18): 14151–14162.
18. Tsiftsis T A. Performance of heterodyne Optical Wireless Communication systems over gamma-gamma atmospheric turbulence channels [J]. Electronics Letters, 2008, 44(5): 373–374.
19. Sandalidis H G, Tsiftsis T A, Karagiannidis G K. Optical wireless communications with heterodyne detection over turbulence channels with pointing errors [J]. Journal of Lightwave Technology, 2009, 27(20): 4440–4445.
20. Djerroud K, Acef O, Clairon A, et al. Coherent optical link through the turbulent atmosphere [J]. Optics letters, 2010, 35(9): 1479–1481.
21. Salem M, Rolland J P. Heterodyne efficiency of a detection system for partially coherent beams [J]. Journal of the Optical Society of America A, 2010, 27(5): 1111–1121.
22. Fice M J, Chiuchiarelli A, Ciaramella E, et al. Homodyne coherent optical receiver using an optical injection phase-lock loop [J]. Journal of Lightwave Technology, 2011, 29(8): 1152–1164.

23. Bayaki E, Schober R. Performance and design of coherent and differential space-time coded FSO systems [J]. Journal of Lightwave Technology, 2012, 30(11): 1569–1577.
24. Kikuchi K. Fundamentals of coherent optical fiber communications [J]. Journal of Lightwave Technology, 2016, 34(1): 157–179.
25. Kakarla R, Mazur M, Schroder J, et al. Power efficient communications employing phase sensitive pre-amplified receiver [J]. IEEE Photonics Technology Letters, 2022, 34(1): 3–6.
26. Sun Yiyi. Laser heterodyne detection in turbulent atmosphere [J]. Chinese Journal of Lasers, 1985, 13(3): 174–177.
27. Qian Xinzhong. Effects of Local Power on Coherent Optical Communication Systems [J]. Journal of Nanjing University of Posts and Telecommunications, 1988, 8(1): 129–135.
28. Zhang Chaoying, Lin Zongqi. Optimal Local Oscillator Power in Optical Heterodyne System [J]. Chinese Journal of Quantum Electronics, 1988, 5(4): 368–369.
29. Zhang X, Jiang P. Limiter-discriminator detection of a coherent optical heterodyne M-ary CPFSK receiver [J]. Journal of Lightwave Technology, 1992, 10(8): 1127–1131.
30. Min Shuangquan, Huang Shuangquan. Experimental Investigation on Endless Polarization Controller [J]. Chinese Journal of Scientific Instrument, 1996, 17(4): 357–360.
31. He Yi. Optical System of an Imagine CO_2 Laser Radar [J]. Acta Optica Sinica, 1999, 19(8): 1079–1083.
32. Wang Qi, Wang Chunhui, Shang Tieliang. Heterodyne Efficiency of Coherent Detection with Guassian Local-Oscillator and Airy Spot Signal Beam [J]. Chinese Journal of Lasers, 2003, 30: 183–186.
33. Xiao Shilin, Zeng Qingji. Applications of Optical Coherent Detection [J]. Optical Coherent Technology, 2003, 11(09): 28–31.
34. Li Weiwen, Zhang Xianmin, Chen Kangsheng, et al. Application of Simulated Annealing Algorithm on Endless Polarization Controller [J]. Acta Photonica Sinica, 2005, 34(6): 820–828.
35. Li Weiwen, Jin Xiaofeng, Zhang Xianmin, et al. Application of genetic algorithms in polarization control [J]. Journal of Zhejiang University(Engineering Scienice), 2006, 40(3): 443–447.
36. Li Cheng, Mao Hongmin, Zhen Shenglai, et al. Influence of Atmospheric Turbulence on Laser Heterodyne Detection [J]. Journal of Atmospheric and Environmental Optics, 2007, 2(4): 247–250.
37. Xu Jing, Mao Hongmin, Zhen Shenglai, et al. The Research for the Deterioration of Spatial Coherence of Laser Heterodyne Detection Caused by Atmospheric Turbulence [J]. Laser & Infrared, 2007, 37(12): 1245–1249.
38. Zhang Wenrui, Zeng Xiaodong, Man Xiangkun. Study on optical heterodyne detection [J]. Infrared and Laser Engineering, 2008, 37(supplement):146–148.
39. Yao Zhoushi, Deng Ke, Li Xiaofeng. The effect of pointing error on performance of satellite coherent laser communication link [J]. Journal of Optoelectronics Laser, 2010, 21(6): 868–870.
40. Gao Long, Wang Chunhui, Li Yanchao, et al. Investigation of Balanced Detection and Receiver for 1.55μm Coherent Wind Lidar [J]. Acta Photonica Sinica, 2010, 39(6): 1064–1069.
41. Liu Hongzhan, Ji Yuefeng, Xu Nan, et al. Effect of Amplitude Profile Difference of Signal and Local Oscillator Wave on Heterodyne Efficiency in the Inter-Satellite Coherent Optical Communication System [J]. Acta Optica Sinica, 2011, 31(10): 1006001–1006006.
42. Xiao Xi, Chen Mingsong. Design of heterodyne asynchronous demodulation receiver for coherent optical communication [J]. Da Zhong Ke Ji, 2011, 140(4): 21–26.
43. Zhang Tong, Tong Shoufeng. Effect of Collimating Mismatch on Mixing Efficiency for Space Coherent Laser Communication [J]. Journal of Changchun University of Science and Technology(Natural Science Edition), 2013, 36(1–2): 13–15.

44. Wu Yunyun, Li Xinyang, Rao Changhui. Effect of Atmospheric Turbulence Aberration on the Bit-Error Performance of Homodyne Binary Phase Shift Keying Coherent Optical Communication [J]. Acta Optica Sinica, 2013, 33(6): 0606002.1–0606002.7.
45. Li Cheng Qiang, Wang Ting Feng, Zhang HeYong, et al. Effect of source parameters on polarization characteristics of electromagnetic beam propagating in atmospheric turbulence [J]. Acta Phys. Sin, 2014, 63(10): 104201.1–104201.8.
46. Liu Hongyang, Zhang Yange, Ai Yong, et al. Design and Implementation of Balance Detector Used in Coherent Optical Communication System [J]. Laser & Optoelectronics Progress, 2014, 51(7): 23–29.
47. Li Xiangyang, Ma Zongfeng, Shi Dele. Effect of Gaussian fields distribution on mixing efficiency for coherent detection [J]. Infrared and Laser Engineering, 2015, 44(2): 539–543.
48. Dai Yonghong, Liu Yanfei, Zhou Haotian, et al. Sensitivity testing and experiment of balanced photodetectors in the space coherent optical communication [J]. Infrared and Laser Engineering, 2015, 44(10): 3110–3118.
49. Chang Shuai, Tong Shoufeng, Jiang Huilin, et al. Optical Phase-Locked Loop Technology in Inter-Satellite High-Speed Coherent Laser Communication Systems [J]. Acta Optica Sinica, 2017, 37(2): 0206004, 1–0206004.9.
50. Wang Mengmeng, Dong Ruifang, Xiang Xiao, et al. Theoretical research for absolute distance measurement based on heterodyne detection principle [J]. Chinese Journal of Scientific Instrument, 2016, 37(8): 1861–1868.
51. Ma Tingting, Tong Shoufeng, Nan Hang, et al. Effect of Polarization Property of Signal Light on Mixing Efficiency of Space Coherent Detection [J]. Laser & Optoelectronics Progress, 2017, 54(2): 020604.1–020604.7.
52. Nan Hang, Zhang Peng, Tong Shoufeng, et al. Analysis of the effect of light spot size and ray axis deflection on heterodyne efficiency of space optical hybrid [J]. Infrared and Laser Engineering, 2017, 46(4): 0422001-1–0422001-8.
53. Lei Fang, Wang Lian. Optical communication system based on the principle of optical coherence detection [J]. LASER JOURNAL, 2018, 39(9): 138–142.
54. Feng Litian, Zhao Pei'e, Shi Xiaoding, et al. Influence of light source linewidth on detection performance of coherent wind lidar [J]. Infrared and Laser Engineering, 2019, 48(4): 0406005.1–0406005.6.
55. Sun Jing, Huang Puming, Yao Zhoushi. Performance of Satellite-to-Ground Laser Communications Under the Influence of Atmospheric Turbulence and Platform Micro-Vibration [J]. Laser & Optoelectronics Progress, 2021, 58(3): 0301003-1–0301003-6.
56. Chen Jinni. Study on key Technologies of subcarrier modullated non-optical heterodyne detection system [D]. Xi'an: Xi'an University of Technology, 2018.
57. Ke X Z, Liang J Y, Xu D S, Wang J F. Research progress of pulse position modulation technology in optical wireless communication [J]. Opto-Electronic Engineering, 2022, 49(3): 3–24.
58. Ke X Z, Wu J L, Yang S J. Research progress and prospect of atmospheric turbulence for Optical Wireless Communication [J]. CHINESE JOURNAL OF RADIO SCIENCE, 2021, 36(3): 323–339.
59. Tan Z K, Ke X Z. Mixing Efficiency in Coherent Detection System [J]. Laser & Optoelectronics Progress, 2017, 54(10): 100605-1–100605-9.
60. Ke X Z, Tan Z K. Effect of angle-of-arrival fluctuation on heterodyne detection in slant atmospheric turbulence [J]. Applied Optics, 2018, 57(5): 1083–1090.
61. Tan Zhenkun. Effects of factors on the performance of heterodyne detection and experimental investigation in Optical Wireless Communication [D]. Xi'an: Xi'an University of Technology, 2021.
62. Ke X Z, Yang S J, Wu J L, et al. Research progress of wireless optical MIMO technology and space-time coding [J]. Journal of Hunan University of Science and Engineering, 2021, 42(5): 25–32.

63. Ke X Z, Yang S J, Wu J L, et al. Research progress of adaptive optics in Optical Wireless Communication system for Xi'an University of Technology [J]. HIGH POWER LASER AND PARTICLE BEAMS, 2021, 33(8): 30–52.
64. Ke X Z, Wu P F. Theory of adaptive optics and its application in Optical Wireless Communication [M]. Beijing: Springer Press, 2024.
65. Ke X Z, Qin H H, Yang S J, et al. Night background light noise model of visible light communication system in vehicle networking environment [J]. CHINESE JOURNAL OF RADIO SCIENCE, 2021, 36(6): 986–990.
66. Chen Mu. Research on channel on noise in wireless optical coherence communication and its effect on coherence detection performance [D]. Xi'an: Xi'an University of Technology, 2018.
67. Ma B B. Experimental study on polarization control in coherent optical communication systems [D]. Xi'an: Xi'an University of Technology, 2021.
68. Ma B B, Ke X Z, Zhang Y. Polarization Control and Control Algorithm of Beams in Coherent Optical Communication System [J]. Chinese Journal of Lasers, 2019, 46(1): 0106002.1–0106002.8.
69. Kong Y X, Ke X Z, Yang Y. Impact of local oscillator power on SNR in space coherent optical communications [J]. Infrared and Laser Engineering, 2016, 45(2): 0222002.1–0222002.6.
70. Kong Y X, Ke X Z, Yang Y. Bit Error Rate of Laser Linewidth in Spatial Coherent Optical Communication Link [J]. Laser & Optoelectronics Progress, 2018, 55(4): 040603.1–040603.6.
71. Kong Y X. Wavefront Correction Technique of Spatial Coherent Optical Communication with LC-SLM [D]. Xi'an: Xi'an University of Technology, 2021.
72. Ke X Z, Chen J N. Experimental Investigation on Non-optical Heterodyne Detection Technology of 1 km Atmospheric Laser Communication System [J]. JOURNAL OF APPLIED SCIENCES—Electronics and Information Engineering, 2014, 32(4): 379–384.
73. Kong Y X, Ke X Z, Yang Y. Influence Research of Atmospheric Turbulence on Space Coherent Optical Communications [J]. Laser & Optoelectronics Progress, 2015, 52(8): 080601.1–080601.7.
74. Yang S J, Ke X Z. Carrier Frequency Stability Control in Coherent Optical Communication [J]. Laser & Optoelectronics Progress, 2018, 55(4): 040601.1–040601.7.
75. Wu jiali. Research of the coherent optical communication system with wavefront sensorless [D]. Xi'an: Xi'an University of Technology, 2018.
76. Mao Yanzi. A Study of mode degeneration and polarization characteristics in coherent optical communication [D]. Xi'an: Xi'an University of Technology, 2018.
77. Wang Yi. Research on DPSK demodulation technology in coherent optical communication system [D]. Xi'an: Xi'an University of Technology, 2021.
78. Bai Jiajun. Research and Implementation of OPSK modulation and demodulation in coherent optical communication system [D]. Xi'an: Xi'an University of Technology, 2024.
79. Ke X Z, Deng Lijun. Partial coherent light transmission theory in optical wireless communication [M]. Beijing: Science Press, 2018.
80. Kong Meng, Lu Yanting, Lin Dong, et al. Digital Laser Frequency Stabilization with Reference to An Optical Frequency Comb [J]. Acta Optica Sinica, 2021, 41(16): 1614001-1–1614001-6.
81. Jiang Wenhan. Overview of adaptive optics development [J]. Opto-Electronic Engineering, 2018, 45(3): 1–15.
82. Babcock H W. The possibility of compensating astronomical seeing [J]. Publications of the Astronomical Society of the Pacific, 1953, 65(386): 229–236.
83. Guan Baozhu, Chen Haiqing, Ye Jiaxiong, et al. Model of beacon used in the air-borne adaptive optical system [J]. Laser Technology, 2004, 28(6): 613–615.
84. Devaney N, Goncharov A V, Dainty J C. Chromatic effects of the atmosphere on astronomical adaptive optics [J]. Applied optics, 2008, 47(8): 1072–1081.
85. Fried D L. Branch point problem in adaptive optics [J]. JOSA A, 1998, 15(10): 2759–2768.

86. Toselli I, Andrews L C. Phillips R L, et al. Angle of arrival fluctuations for free space laser beam propagation through non kolmogorov turbulence [J]. ATMOSPHERIC PROPAGATION IV, 2007, 65(1):149–160.
87. David V, Erandi W, Andreas M, et al. Computer simulations of optical turbulence in the weak- and strong-scattering regime: angle-of-arrival fluctuations obtained from ray optics and wave optics [J]. Optical Engineering, 2018, 57(10):102–104.
88. Conan R, Borgnino J, Ziad A, et al. Analytical solution for the covariance and for the decorrelation time of the angle of arrival of a wave front corrugated by atmospheric turbulence [J]. JOSA A, 2000, 17(10): 1807–1818.
89. Ben-Yosef N, Goldner E, Weitz A. Two-color correlation of scintillations [J]. Applied optics, 1986, 25(19): 3486–3489.
90. Gorelaya A V, Shubenkova E V, Dmitriev D I, et al. Investigation of dual-wavelength laser beam propagation along the in-door atmospheric path[C]//Optics in Atmospheric Propagation and Adaptive Systems XVIII. SPIE, 2015, 9641: 88–93.
91. Li F, Rong J, Zhong X C, et al. Study on correcting angle-of-arrival fluctuations of space optical communication with AO [J]. SECOND INTERNATIONAL CONFERENCE ON SPACE INFORMATION TECHNOLOGY, 2007, 67(9):141–146.
92. Kibblewhite E J, Chun M R. Design of tip-tilt and adaptive optics servos using measured angle-of-arrival and phase power spectra[C]//Adaptive Optical System Technologies. SPIE, 1998, 3353: 522–530.
93. Lukin V P. Efficiency of some correction systems [J]. Optics Letters, 1979, 4(1): 15–19.
94. Jolissaint L, Kendrew S. Modeling the Chromatic Correction Error in Adaptive Optics: Application to the Case of Mid-Infrared Observations in Dry to Wet Atmospheric Conditions [C]. Adaptative Optics for Extremely Large Telescopes. EDP Sciences: 2010, 05021:1–5
95. Ke X Z, Chen X Z. Correcting wavefront distortion of dual-wavelength beams due to atmospheric turbulence with a correction coefficient [J]. Optics and Photonics Journal, 2020, 10(4): 64–77.
96. Zhang Yixin. Arrival Angle of Beam Wave Propagation in Turbulent Atmosphere [J]. Chinese Journal of Quantum Electronics, 1987, 12(1): 70–75.
97. Gu Demen, Qin Kecheng. Statistical optics [M]. Beijing: Science Press, 1992.
98. Wang Zhuxi. Introduction to special functions [M]. Beijing: Science Press, 1979.
99. Ke X Z, Zhang D Y. Fuzzy control algorithm for adaptive optical system [J]. Applied Optics, 2019, 58(36):9967–9975.
100. Chen Chunyi, Yang Huamin, Jiang Huilin, et al. Performance Analysis and Optimization of Partially Coherent Optical Communication Links Through Atmospheric Channel [J]. Acta Electronica Sinica, 2009, 30(8): 1869–1872.
101. Abbaas G L, Chan V, Yee T. A dual-detector optical heterodyne receiver for local oscillator noise suppression [J]. Journal of Lightwave Technology, 1984, 3(5): 1110–1126.
102. Slusher R E, Hollberg L W, Yurke B, et al. Observation of squeezed states generated by four-wave mixing in an optical cavity [J]. Physical Review Letters, 1985, 55(22): 2409–2412.
103. He Ning, Feng Taiqin, Liao Xin. Research on Noise Cancellation Based on Acousto-Optic [J]. Acta Optica Sinica, 2015, 35(07): 78–85.
104. Nan Hang, Zhang Peng, Tong Shoufeng, et al. Performance Analysis of Free Coherent Optical Communication in Atmosphere Turbulence with Tracking Error [J]. Acta Photonica Sinica, 2015, 44(08): 56–61.
105. Liang Hexi, Dai Yonghong, Ai Yong, et al. Influence of local oscillator power on sensitivity of coherent detection of space balance detector [J]. Optics and Precision Engineering, 2017, 25(02): 334–341.
106. Anaka T, Taguchi M, Tanaka K. Heterodyne efficiency for a partially coherent optical signal [J]. Applied Optics, 1992, 31(25): 5391–5394.
107. Pu Lili, Zhou Yu, et al. Receiving Characteristics of Coherent Ladar under Partially Coherent Condition [J]. Acta Optica Sinica, 2011, 31(12): 268–276.

108. Salem M, Rolland J P. Heterodyne efficiency of a detection system for partially coherent beams [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2010, 27(5): 1111–1121.
109. Li C Q, Wang T F, Zhang H Y, et al. The performance of heterodyne detection system for partially coherent beams in turbulent atmosphere [J]. Optics Communications, 2015, 356(23): 620–631.
110. Salem M, Rolland J P. Effects of coherence and polarization changes on the heterodyne detection of stochastic beams propagating in free space [J]. North-Holland, 2008, 281(20): 5083–5091.
111. Wang Y, Li C Q, Wang T F, et al. The effects of polarization changes of stochastic electromagnetic beams on heterodyne detection in turbulence [J]. Laser Physics Letters, 2016, 13(11), 116006.
112. Wolf E. Unified theory of coherence and polarization of random electromagnetic beams [J]. Physics Letters A, 2003, 312: 263267.
113. Salem M, Dogariu A. Optical heterodyne detection of random electromagnetic beams [J]. Optica Acta International Journal of Optics, 2004, 51(15): 2305–2313.
114. Gori F, Santarsiero M, Piquero G, et al. Partially polarized Gaussian Schell-model beams [J]. Journal of Optics A: Pure and Applied Optics, 2001, 3(1): 1–9.
115. Li C Q, Zhang H Y, Wang T F, et al. Investigation on coherence characteristics of Gauss-Schell model beam propagating in atmospheric turbulence. Acta Phys. Sin.[J]. Acta Physica Sinica, 2013, 62(22): 1–6.
116. Duan M L, Li J H, Wei J L. Influence of different propagation paths on the propagation of laser in atmospheric turbulence [J]. Optoelectronics Letters, 2013, 9(6): 477–480.
117. Zhai C, Tan L Y, Yu S Y, et al. Fiber coupling efficiency for a Gaussian-beam wave propagating through non-Kolmogorov turbulence. Optics Express. 2015, 23(12): 15242–15255.
118. ITU-R Document 3J/31-E. On propagation data and prediction methods required for the design of space-to-earth and earth-to-space optical communication systems [C]. Proceedings of the Radio Communication Study Group Meeting (International Telecommunication Union), Budapest, 2001, 277–293.
119. Liu C, Chen C Q, Li X Y, et al. Performance evaluation of adaptive optics for atmospheric coherent laser communications [J]. Optics Express, 2014, 22(13): 15554–15563.
120. Ma Zongfeng, Zhang Chunxi, Zhang Zhaoyang, et al. Signal-Noise Ratio in Optical Heterodyne Detection [J]. Acta Optica Sinica, 2007, 28(05): 889–892.
121. Fink, David. Coherent Detection Signal-to-Noise [J]. Applied Optics, 1975, 14(3):689–690.
122. Tanaka K, Ohta N. Effects of tilt and offset of signal field on heterodyne efficiency [J]. Applied Optics, 1987, 26(4): 627–32.
123. Xiang Jingsong, Pan Lechun. Heterodyne Efficiency and the Effects of Aberration for Space Coherent Optical Communication [J]. Opto-Electronic Engineering, 2009, 36(11): 53–57.
124. Liu Hongzhan, Ji Yuefeng, Xu Nan, et al. Effect of Amplitude Profile Difference of Signal and Local Oscillator Wave on Heterodyne Efficiency in the Inter-Satellite Coherent Optical Communication System [J]. Acta Optica Sinica, 2011, 31(10): 71–76.
125. Li Xiang, Yang Jinhua, Jiang Chengha, et al. Study on collimation mismatch in laser heterodyne detection [J]. Laser & Infrared, 2014, 44(3): 250–253.
126. Tang Ming, Li Weihao, Zhang Mingming, et al. Self-Homodyne Coherent Optical Transmission Techniques [J]. Chinese Journal of Lasers, 2022, 49(12): 223–237.
127. Li G. Recent advances in coherent optical communication [J]. Advances in Optics and Photonics, 2009, 1(2): 279–307.
128. Li C, Wang Y, Wang J, et al. Convolutional Neural Network-Aided DP-64 QAM Coherent Optical Communication Systems [J]. Journal of Lightwave Technology, 2022, 40(9): 2880–2889.
129. Liu W, Jin D, Shi W, et al. Performance analysis of coherent optical communication based on hybrid algorithm [J]. Optics & Laser Technology, 2022, 149: 107878.

130. Modeling the Heterodyne Efficiency of Array Detector Systems in the Presence of Target Speckle [J]. IEEE Photonics Journal, 2019, 11(4):1–9.
131. Kong Yingxiu, Ke Xizheng, Yang Yuan. Influence Research of Atmospheric Turbulence on Space Coherent Optical Communications [J]. Laser & Optoelectronics Progress, 2015, 52(08): 95–101.
132. Li Yun, Li Zhengxuan, Huang Xingang, et al. Coherent Detection System Based on Phase-Modulated Local Oscillator [J]. Acta Optica Sinica, 2021, 41(20):15–23
133. Guo Z. Analysis of signal-to-noise ratio and heterodyne efficiency for reference-beam laser Doppler velocimeter [J]. Optics & Laser Technology, 2012, 44(1):108–113.
134. Zhu X, Kahn J M. Free-space optical communication through atmospheric turbulence channels [J]. IEEE Transactions on communications, 2002, 50(8): 1293–1300.
135. Dai Zhengshuang, Jiang Lun, Song Zhihua, et al. Coherent Detection Mixing Systems with Shaping Elements [J]. Laser & Optoelectronics Progress, 2021, 58(7): 132–138.
136. Chen R S, Zhang X Q, Wang J H, et al. Scalable detection of photonic topological charge using radial phase grating [J]. Applied Physics Letters, 2018, 112(12): 122602.
137. Guo Daifang, Zhang Peng, Gong Xiyu, et al. Design and Simulation Analysis of Laguerre Gaussian Mode Demultiplexing Hybrid Based on Multi-Plane Light Conversion [J]. Chinese Journal of Lasers, 2022, 49(9): 127–143.
138. Wen H, Liu H, Zhang Y, et al. Scalable Hermite–Gaussian mode-demultiplexing hybrids [J]. Optics Letters, 2020, 45(8): 2219–2226.
139. Xie G, Li L, Ren Y, et al. Performance metrics and design considerations for a free-space optical orbital-angular-momentum–multiplexed communication link [J]. Optica, 2015, 2(4): 357–365.
140. Hen Y, Wang X, Xie Z, et al. Optical vortices 30 years on: OAM manipulation from topological charge to multiple singularities [J]. Light: Science & Applications, 2019, 8(1): 1–29.
141. Ren Y, Xie G, Huang H, et al. Adaptive-optics-based simultaneous pre-and post-turbulence compensation of multiple orbital-angular-momentum beams in a bidirectional free-space optical link [J]. Optica, 2014, 1(6): 376–382.
142. Yoshida Y, Umezawa T, Kanno A, et al. A Phase-Retrieving Coherent Receiver Based on Two-Dimensional Photodetector Array [J]. Journal of Lightwave Technology, 2020, 38(1): 90–100.
143. Noll R J. Zernike Polynomials and Atmospheric Turbulence [J]. Journal of the Optical Society of America, 1976, 66(3): 207–211.
144. Andrade P P, Garcia P J V, Correia C M, et al. Estimation of atmospheric turbulence parameters from Shack-Hartmann wavefront sensor measurements [J]. Monthly Notices of the Royal Astronomical Society, 2019, 483(1): 1192–1201.
145. Tyson R K, Canning D E. Indirect measurement of a laser communications bit-error-rate reduction with low-order adaptive optics [J]. Applied Optics, 2003, 42(21): 4239–4243.
146. Banakh V A, Zhmylevskii V V, Ignatiev A B, et al. Optical-beam wavefront control based on the atmospheric backscatter signal [J]. Quantum Electronics, 2015, 45(2): 153–160.
147. Betanzos-Torres M A, Castillo-Mixcóatl J, Muoz-Aguirre S, et al. Adaptive optics system simulator [J]. Optics & Laser Technology, 2018, 105(10): 118–133.
148. Norris B R M, Wei J, Betters C H, et al. An all-photonic focal-plane wavefront sensor [J]. Nature Communications, 2020, 11(1): 1–9.
149. Quintavalla M, Bergomi M, Magrin D, et al. Correction of non-common path aberrations in pyramid wavefront sensors to recover the optimal magnitude gain using a deformable lens [J]. Applied Optics, 2020, 59(17): 5151–5157.
150. Yang Y, Yan C, Hu C, et al. Modified heterodyne efficiency for coherent laser communication in the presence of polarization aberrations [J]. Optics Express, 2017, 25(7): 7567–7591.
151. Stotts L B, Andrews L C. Adaptive optics model characterizing turbulence mitigation for free space optical communications link budgets [J]. Optics Express, 2021, 29(13): 20307–20324.

152. Farrell T C. Fast simulation of Strehl loss due to phase aberration for the sizing of adaptive optics in laser communications system design [J]. Applied Optics, 2014, 53(1): 64–70.
153. Ma Chun bo, Shi Jun jie, Wang Ying, et al. An Adaptive Photoelectric Array Signal Processing Algorithm for Free Space Optical Communication [J]. Acta Electronica Sinica, 2021, 49(10): 1908–1912.
154. Xu Dong ling, Yue Peng, Yi Xiang. A Theoretical Analysis Method of OAM-Based FSO Error Performance [J]. Acta Electronica Sinica, 2021, 49(10): 1934–1944.
155. Liu Xi guo, Liu Min, Mao Zhong yang, et al. Analysis of Geometric Attenuation Model of Gaussian Beam under Pointing Error [J]. Acta Electronica Sinica, 2021, 49(10): 1893–1899.
156. Ren J Y, Sun H Y, Zhang L X, et al. Receiving characteristics analysis of large dynamic range based on four-optic coherent mixing [J]. Optik, 2020, 221: 1–7.
157. Chen Dan, Lei Yu, Ke Xi zheng. Performance Analysis of Wireless Optical Adaptive Subcarrier MDPSK Modulation System [J]. Acta Electronica Sinica, 2018, 46(7): 1748–1753.
158. Magidi S, Jabeena A. Free Space Optics, Channel Models and Hybrid Modulation Schemes: A Review [J]. Wireless Personal Communications, 2021, 119(4): 2951–2974.
159. Liu C, Chen S, Li X, et al. Performance evaluation of adaptive optics for atmospheric coherent laser communications [J]. Optics Express, 2014, 22(13): 15554–15563.
160. Geng C, Li F, Zuo J, et al. Fiber laser transceiving and wavefront aberration mitigation with adaptive distributed aperture array for free-space optical communications [J]. Optics Letters, 2020, 45(7): 1906–1909.
161. Konwar S, Boruah B R. Leveraging the orthogonality of Zernike modes for robust free-space optical communication [J]. Communications Physics, 2020, 3(1): 1–9.

Factors Affecting Coherent-Detection Sensitivity

22

Contents

Introduction 800
Direct Detection and Coherent Detection 801
Direct Detection 801
Coherent Detection 802
Signal-to-Noise Ratio and Sensitivity of Coherent-Detection Systems 803
Foreign Research Progress 805
Domestic Research Progress 809
Factors Affecting the Coherent-Detection Limit 813
Thermal and Shot Noise 814
Wavefront Distortion of the Signal Light 815
Polarization Mismatch Between the Signal Light and Local-Oscillator Light 816
Spot-Size Deviation and Optical-Axis Deflection 817
Coupling Efficiency 818
Detector Performance 820
Preoptical Amplifier 821
Development Trends 822
High-Detection Sensitivity and Resolution 822
Wider Applications 822
Simpler and More Efficient Detection Methods 823
Further Research 823
Future Prospects 823
References 824

Abstract

Coherent detection plays an important role in coherent optical communication. Highly sensitive coherent-detection systems can significantly increase the distance of coherent optical communication. Targeted corrections can be made to improve the detection sensitivity by studying the factors that affect it. The principle of coherent detection is to coherently mix the signal light and local-oscillator light and then recover the signal, according to the information carried by the mixed signal. This chapter analyzes the influence of thermal noise, shot

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_22

noise, wavefront distortion, polarization misalignment, beam-spot size deviation, optical-axis deflection, coupling efficiency, detector performance, preoptical amplifiers, and other factors on the sensitivity of coherent detection. We summarize the research results of domestic and foreign scholars in this field in recent years and, finally, provide prospects for its future development direction.

Keywords

Coherent detection · Detection limit · Detection sensitivity · Mixing efficiency · Atmospheric turbulence

Introduction

In 1955, Forrester et al. caused two beams of light with different wavelengths to interfere with each other and generated a new intermediate-frequency signal [1, 2], introducing coherent detection to the world. Coherent-detection methods can be divided into heterodyne, homodyne, and endodyne, according to the detection method.

When the angular frequency of the signal-light center is the same as that of the local-oscillator (LO) light center, it is called homodyne detection. When the central angular frequency of the signal light is different from that of the LO light, it is called heterodyne detection. The signal light and the local-oscillator light are interfered and mixed by a 90° mixer to obtain two orthogonal phase signals; then, the two signals are detected, which is called internal-difference detection.

Coherent-detection sensitivity refers to the limit of what a detector can detect at a certain signal-to-noise ratio or bit-error rate. Theoretically, the detection limit can reach one photon; however, in practical situations, this limit cannot be reached, owing to atmospheric turbulence, thermal noise of the detector, low mixing efficiency, and low detector sensitivity.

Compared with a direct-detection system, the detection sensitivity of a coherent optical communication system can have a gain of approximately 20 dB [3]. It has many modulation methods [4], excellent selectivity [5], large channel capacity [6], and other characteristics that can greatly improve the performance of optical wireless communication systems.

In subsequent research, domestic and foreign scholars have analyzed the influence of thermal noise, shot noise, wavefront distortion, polarization misalignment, spot-size deviation, optical-axis deflection, coupling efficiency, detector performance, preoptical amplifiers, and other aspects of coherent-detection systems on the coherent-detection limit. After years of research, it has become possible to use wavefront-correction systems and beam-polarization control technology. A series of methods, such as fiber coupling and beam control, that reduce the noise interference, improve the detector performance, and amplify the signal can be used to improve the sensitivity of the detection system.

The coherent-detection limit directly affects the performance of coherent communication systems. As the communication distance increases, the signal intensity that can be detected by the receiving end is reduced, and the sensitivity requirements for coherent detection increase. Therefore, it is necessary to conduct a detailed analysis of the factors affecting the coherent-detection limit, which is essential for improving the performance of coherent-detection systems.

This chapter introduces the research progress in coherent detection at home and abroad, as well as research on detection limits by domestic and foreign scholars. The factors that affect the detection limit are analyzed and the factors that suppress these effects are summarized. Finally, the development direction of the coherent-detection limit is summarized and prospected.

Direct Detection and Coherent Detection

Direct Detection

Direct detection refers to directly projecting a signal light onto a photosensitive surface; then, the photodetector outputs the corresponding current based on the intensity of the signal light. As shown in Fig. 22.1, when the signal light reaches the receiving end, the aperture diaphragm and focusing lens converge the light signal onto the photosensitive surface of the photodetector, and the photodetector converts the light intensity of the signal into a voltage or current signal.

Direct detection has a high reliability, simple implementation, and low cost; however, its detection sensitivity is too low to meet the requirements of high-sensitivity detection in long-distance optical wireless communication [7].

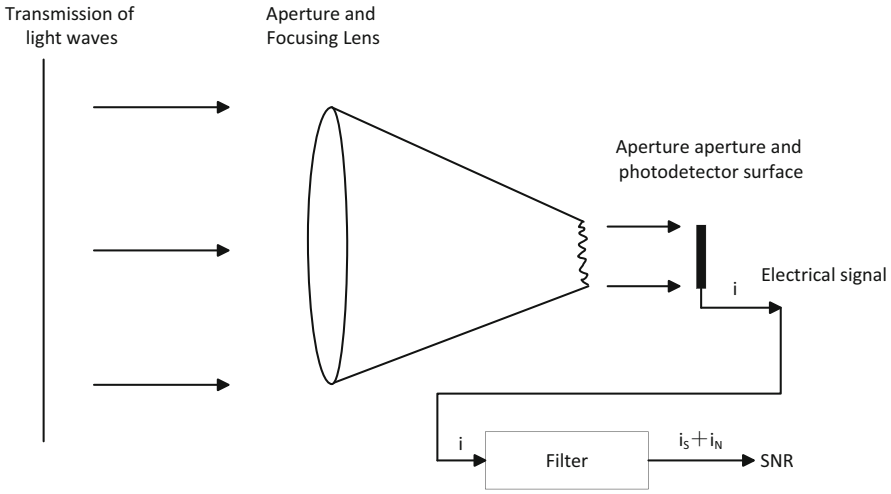


Fig. 22.1 Direct-detection system diagram model [7]

Coherent Detection

Principle of Coherent Detection

Coherent detection involves interference between the local-oscillator light and the signal light. It uses the changes in the light and dark fringes generated by the interference, as well as their intensity and displacement, to determine the coherence and phase difference between the two coherent light beams. This principle is illustrated in Fig. 22.2.

The signal light $E_s(r)$ interferes with the local-oscillator light $E_l(r)$ on the optical mixer, and the interfered light wave is incident on the balanced detector. The balanced detector can detect the interference fringes that form a differential signal. Coherent detection can be divided into zero-difference, internal-difference, and external-difference detection.

Optical mixers are generally composed of fiber couplers and nonlinear optical crystals. The function of a balanced detector is to detect weak optical signals and convert them into electrical signals. The electrical signal-processing module can recover the baseband signal after a series of processing.

Zero-Difference Detection

When the frequencies of the local oscillator and the signal light are the same, it is called homodyne detection. The principle block diagram of zero-difference detection is shown in Fig. 22.3. The signal light and local-oscillator light become electrical signals after demodulation through optical phase-locked coherence. The electrical signal can be obtained as a baseband signal after passing through a low-pass filter, and the original signal can be obtained from the decision output.

An optical phase-locked loop can stabilize the frequencies and phases of the local oscillator and signal light [8, 9]; however, it increases the complexity of the receiver. The advantage of homodyne detection is high sensitivity.

Heterodyne Detection

When the frequencies of the local-oscillator light and signal light are not equal, the coherent-detection method is also called heterodyne detection and can be divided into synchronous and asynchronous demodulation, based on the demodulation method at the receiving end.

A schematic diagram of synchronous demodulation is shown in Fig. 22.4a. The signal light and local-oscillator light form an intermediate signal after passing

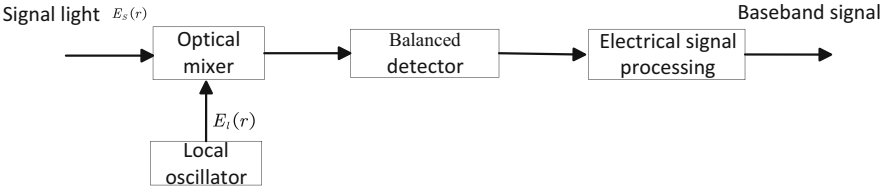


Fig. 22.2 Principle block diagram of a coherent-detection system [3]

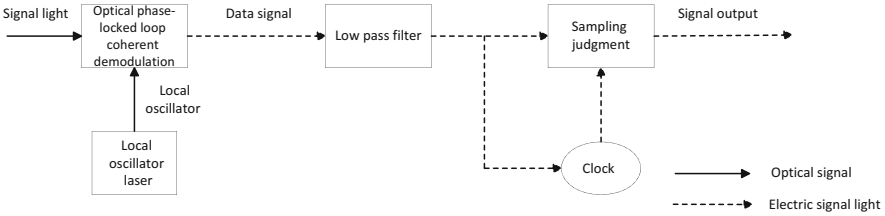


Fig. 22.3 Demodulation-principle block diagram of a zero-difference detection system

through the optical mixer. The signal then passes through the photodetector and becomes an electrical signal. After the carrier frequency is recovered, it is multiplied by the intermediate-frequency signal. The signal output of the multiplier is then passed through a low-pass filter to obtain the baseband signal.

The principle of asynchronous demodulation is shown in Fig. 22.4b. After passing through the mixer and photodetector, the signal light and local-oscillator light pass directly through a bandpass filter and then undergo envelope detection and other processing to obtain the baseband signal directly. Although this method is relatively simple, it has strict requirements for the phase and frequency of the local-oscillator and signal lights.

Internal-Difference Detection

A structural diagram of an internal-difference detection system is shown in Fig. 22.5. The transmitting end is mainly comprised of signal generators, laser sources, and modulators. At the receiving end, the signal light and local-oscillator light pass through a mixer to obtain an intermediate-frequency signal. It passes through a balanced detector, and the electrical signal is then processed by a PC to obtain the original signal.

Signal-to-Noise Ratio and Sensitivity of Coherent-Detection Systems

In coherent-detection systems, the detection sensitivity refers to the limit of what a detector can detect at a certain signal-to-noise ratio (SNR) or bit-error rate (BER). In theory, the detection limit of an analyzer can reach one photon; however, in reality, the detection limit cannot be reached because of atmospheric turbulence, thermal noise of the detector, low mixing efficiency, and low detector sensitivity. This section analyzes the relationship between the sensitivity and signal-to-noise ratio of coherent-detection systems.

The signal-to-noise ratio of a photodetector is defined as follows [3]:

$$\mathrm{SNR}_{\mathrm{IF}} = \frac{\left\langle i_{\mathrm{IF}}^2 \right\rangle}{\left\langle i_{\mathrm{N}}^2 \right\rangle} = \frac{\eta \left| \int_S E_{\mathrm{l}} \cdot E_{\mathrm{s}}^* \, \mathrm{d}s \right|^2}{h\nu \Delta f_{\mathrm{IF}} \cdot \int_S A_l^2 \, \mathrm{d}s} = \frac{\eta \left| \int_S E_{\mathrm{l}} \cdot E_{\mathrm{s}}^* \, \mathrm{d}s \right|^2}{h\nu \Delta f_{\mathrm{IF}} \cdot \int_S E_{\mathrm{l}} \cdot E_{\mathrm{l}}^* \, \mathrm{d}s} \tag{22.1}$$

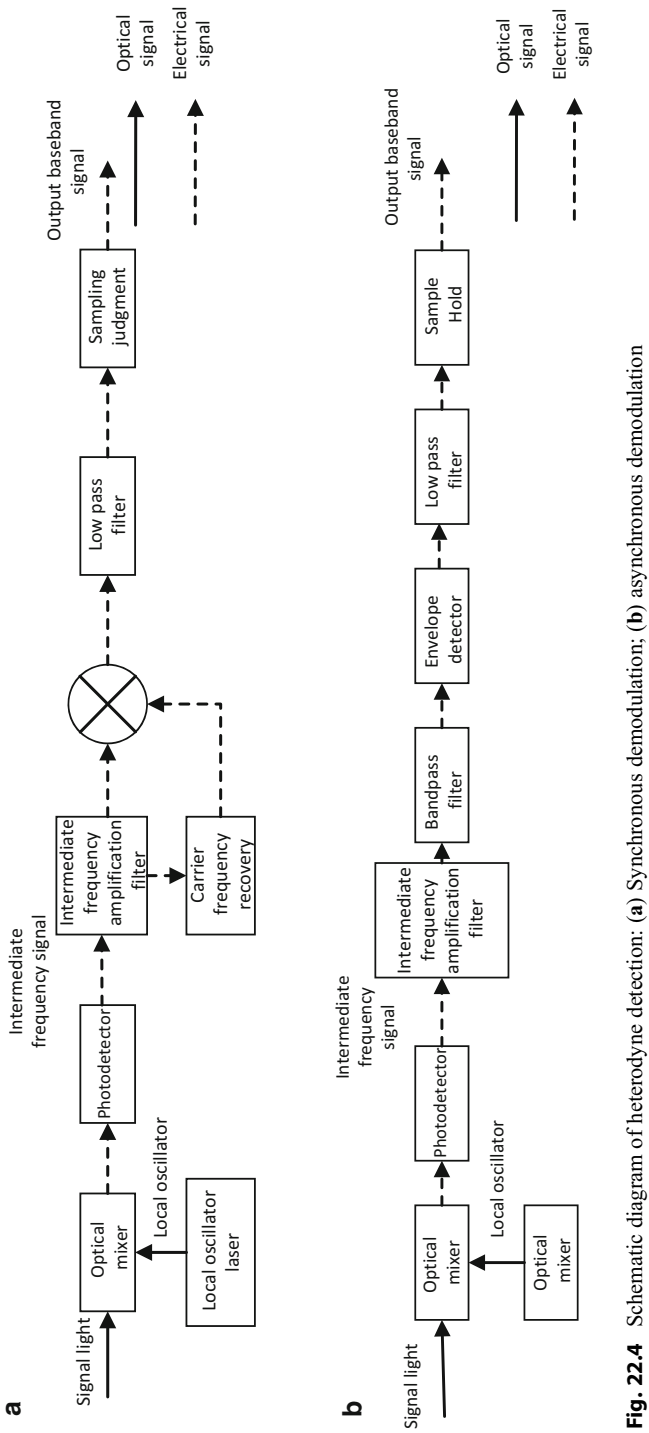


Fig. 22.4 Schematic diagram of heterodyne detection: (**a**) Synchronous demodulation; (**b**) asynchronous demodulation

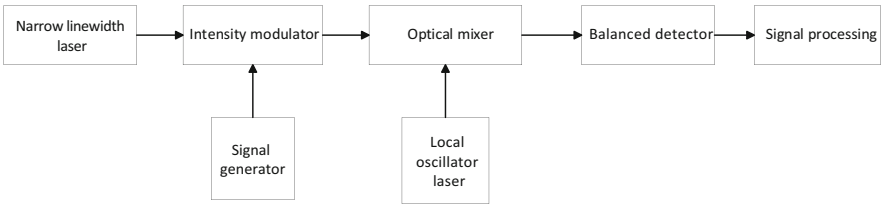


Fig. 22.5 Basic structure of the internal-difference detection system

where $<\cdot>$ is the time average, i_F is the intermediate-frequency current, i_N is the noise current, η is the quantum efficiency of the ideal photosensitive surface, A_L is the amplitude of the local-oscillator light, E_l is the local-oscillator light electric field, E_s is the signal photoelectric field, and Δf_{IF} is the effective noise bandwidth.

When the mixing efficiency reaches the ideal value, the signal-to-noise ratio can be expressed as [3]

$$\mathrm{SNR}_{IF} = \mathrm{SNR}_0 = \frac{\eta P_S}{h\nu\Delta f} = \frac{2\eta P_S}{h\nu B}, \tag{22.2}$$

where η is the quantum efficiency, h is Planck's constant, P_S is the signal optical power, B is twice the effective noise bandwidth, and ν is the carrier frequency.

From Eq. (22.2), it can be seen that when SNR is 1, NEP $= P_S = h\nu\Delta f/\eta$. This variable is called the coherent-detection sensitivity. When the quantum efficiency of the mixer reaches 1, $\Delta f = 1$ Hz, and the ultimate coherent-detection sensitivity can be obtained as one photon. However, in practical coherent detection, such high sensitivity cannot be achieved. After years of research, the detection-sensitivity limit has reached the nanowatt level [3].

Foreign Research Progress

Compared to direct detection, the sensitivity of coherent detection can be significantly improved, while improving the performance of coherent communication systems. This has led many foreign scientists to invest in this field. They have analyzed the impact on coherent-detection limits of aspects such as wavefront distortion, detector sensitivity, intensity noise caused by the local-oscillator light, spot size and optical-axis deflection, coupling efficiency, and polarization mismatch. Methods, such as wavefront correction, polarization control, high-sensitivity detectors, and searching for the optimal local-oscillator power, have been designed to continuously improve the detection sensitivity and approach the coherent-detection limit.

In 1966, Teich et al. [10] measured the signal-to-noise ratio (SNR) and minimum detectable power of m-scattering radiation heterodyne detection. Using photoconductive Ge: Cu as the detector, the minimum detection power observed at a 70 kHz frequency and a 270 kHz bandwidth was 3.5×10^{-14} W. This corresponds to 1.3 at a

bandwidth of 1 MHz. The minimum detectable power of 10^{-19} W is within ten times that of the theoretical photon counter.

In 1972, Popoola et al. [11] developed an experimental device that could compensate for the influence of turbulence in real time. The experimental results showed that the device effectively compensated for the wavefront distortion caused by atmospheric turbulence in a 300 m horizontal link.

In 1975, Cohen et al. [12] thoroughly studied the effects of a series of factors, such as spatial-angle mismatch, spot size, and detector performance, on the mixing efficiency of a mixer. The research results showed that when the signal light had an Airy distribution, the spatial-angle mismatch had a relatively small impact on the mixing efficiency.

In 1983, Abbas et al. [13] calculated the signal-to-noise ratio of homodyne/heterodyne receivers for two detectors. Their results showed that excessively intense noise in a local oscillator could be eliminated without reducing the performance of the receiver. Later, it was experimentally proven that using two matched detectors could achieve near-quantum-limit performance.

In 1987, Tanaka et al. [14] analyzed the impact of the vector-tilt angle and offset error of the signal light on the mixing efficiency. The results showed that when the vector-tilt angle was less than 10^{-4} degrees and the offset reached one-fifth of the incident aperture, the mixing efficiency exceeded 60%.

In 1988, Calvani et al. [15] developed a primitive coherent-transmission technology with polarization modulation and dual-channel differential heterodyne detection. In a preliminary experiment, the detection-sensitivity limit reached 4 dB and showed high immunity to laser phase noise.

In 1992, Tanaka et al. [16] studied the mixing efficiency of a detection system in which the local-oscillator and signal lights were partially coherent Gaussian beams. The results showed that if ratio A is the minimum $1/e$ radius of the signal light to the minimum $1/e$ radius of the local-oscillator, and ratio B is the input aperture to the detector aperture, A and B are inversely proportional.

In 1998, Bar David et al. [17] used mathematical theory to demonstrate the superior performance of coherent dual-light detection over single-diode detection and established the basic principle of local-oscillator intensity. They proposed modulation and detection formats that reduced additional noise using the influence of phase noise.

In 1999, Wang et al. [18] proposed a novel antireception heterodyne-detection scheme for a monolithic integrated four-segment coherent transceiver. The proposed antireception heterodyne-detection scheme further simplified the design of integrated coherent transceivers and promoted the use of a single terminal device for duplex transmission.

In 2000, Sato et al. [19] studied and designed a phase-drift suppression method (PDSM), which suppressed the influence of phase drift and improved the stability of heterodyne beat signals, and demonstrated its application.

In 2001, Gol'tsman et al. [20] developed a new type of high-bandwidth laser detector. The experimental results showed that the detector could achieve long-distance optical-coupling transmissions with a coupling efficiency of 38%. The

measurement bandwidth reached 17.5 MHz and the detection sensitivity reached 0.52 μW.

In 2002, Ricklin et al. [21] demonstrated a light-coupling technology that used integrated circuits to control a deformable mirror for wavefront correction. The results showed that the fiber-coupling efficiency using the new coupling technology reached 75%.

In 2003, Sobolewski et al. [22] proposed a new type of ultrafast single-photon detector, which was made of NbN films with thicknesses of 3.5 nm and 10 nm. For photons at 1550 nm, the detection probability was 3.5%. The response time and jitter of the detector were 100 ps and 35 ps, respectively, which were limited by the acquisition system. At the optimal bias voltage, the dark count was less than 0.01 per second. In terms of the counting rate, jitter, and dark counting, NbN single-photon detectors are superior to semiconductor detectors.

In 2004, Fernandez et al. [23] conducted an experiment on a coherent-optical communication system using a 1064 nm wavelength laser. In the experiment, the laser powers of the local-oscillator and signal lights were 50 mW and 23 mW, respectively. The heterodyne-detection method was used to obtain a frequency-difference signal of 6 MHz from the mixer.

In 2005, Yang et al. [24] studied a radial-shear interference system with spatial-phase modulation that could be used to detect wavefront distortions and control the beam quality. This interferometer was used to test the wavefront distortion of laser pulses with a 1064 nm near-infrared wavelength and 10 ns pulse time width. The software displayed the results of the pulse wavefront quality and energy distribution. The measured beam diameter reached 150 mm, and the root-mean-square test accuracy exceeded 1/15 wavelength.

In 2007, Dauler et al. [25] proposed a multielement superconducting-nanowire single-photon detector that consisted of multiple independently biased elements that formed a continuous active region. Two superconducting-nanowire single-photon detectors with two components were experimentally produced and tested. There was no measurable crosstalk between the devices, and the relative timing jitter was less than 50 ps. This was four times the maximum counting rate of a single superconducting-nanowire single-photon detector with the same active area. Compared with single-element superconducting-nanowire single-photon detectors, multi-element superconducting-nanowire single-photon detectors have larger effective areas and higher speeds.

In 2007, Park et al. [26] developed a waveguide photodetector with a hybrid waveguide structure consisting of AlGaInAs quantum wells on a silicon waveguide. The light in the hybrid waveguide was absorbed by the AlGaInAs quantum wells under a reverse bias. The photodetector had a fiber-coupling responsivity of 0.31 A/W, which was 1.5 μm within the wavelength range of m; it had an internal quantum efficiency of 90%.

In 2009, Xu et al. [27] developed a novel homodyne-detection structure that made orthogonal and sequential in-phase measurements of the signal light. The results showed that the weaker the equivalent substrate, the closer the detection limit to the standard quantum limit.

In 2010, Gregory et al. [28], in view of the impact of wavefront distortion on detection sensitivity, installed a coherent laser-communication terminal on the ground and used the terminal to establish a 5.635-Gbit/s communication link with a satellite. Experimental results showed that the adaptive-optics technology used in the link effectively suppressed the wavefront distortion of the signal light.

In 2010, Salem et al. [29] derived an analytical expression for the heterodyne efficiency when two partially coherent Gaussian–Schell-model beams were coherently mixed on the surface of a photodetector. They provided numerical examples of the heterodyne efficiency changing with the misalignment angle, detector radius, and overlapping beam parameters. The experimental results showed that partially coherent beams were more susceptible to a decrease in the heterodyne efficiency than coherent beams.

In 2013, Gan et al. [30] developed a new type of graphene photodetector, which had the advantages of high responsiveness, high speed, and a wide spectral bandwidth. Using a metal-doped graphene junction coupled with the evanescent waveguide, the detector achieved a light response of more than 0.1 $\mathrm{A\,W^{-1}}$, and a uniform response between 1450 nm and 1590 nm. Under zero-bias operation, it demonstrated a response rate of more than 20 GHz and a 12 Gbit/s optical data link.

In 2014, Boroson et al. [31] implemented a communication link between the ground and the Moon with an information-transmission rate of 622 Mbit/s. The project combined ground-terminal array-transceiver technology with superconducting-nanowire single-photon detection to solve the problems of light-intensity flickering and low receiving power, and to improve the detection sensitivity.

In 2015, Chengqiang et al. [32] introduced the power spectrum of atmospheric turbulence to study the effects of the atmosphere on heterodyne systems. Based on the obtained results, the relationships between the heterodyne efficiency and the detector diameter, misalignment angle, and beam parameters under different turbulence conditions were shown.

In 2016, Wang et al. [33] studied the variation in the heterodyne efficiency with the misalignment angle, detector diameter, turbulence conditions, and overlapping beam parameters based on the Tatarskii spectral model. The results indicated that the variation in the heterodyne efficiency depends on the initial beam polarization, and for larger detector diameters, a turbulent atmosphere significantly reduces the heterodyne efficiency.

In 2017, Yang et al. [34] analyzed and illustrated the effect of polarization aberrations on the output polarization state using polar and Pauli–Zernike decomposition algorithms. Based on the modified heterodyne efficiency, a model of an off-axis optical system with a periscope scanner was established and used to discuss the variation in heterodyne efficiency. The results showed that the improved heterodyne efficiency not only comprehensively described the coherent-detection system but could also be used to evaluate and minimize the polarization aberration of the optical system.

In 2018, Chen et al. [35] developed optical-fiber coupling equipment suitable for long-distance communication, which was used to compensate for random angular

errors and residual aberrations. The experimental results showed that the coupling efficiency reached 73% and the mean-squared error reached 1.1% after using this device. The sensitivity of the coherent-detection system was approximately 20 dB higher than that of the direct-detection system.

In 2020, Maiti et al. [36] studied a MoTe-based photodetector that exhibited a strong optical response in the 1550 nm wavelength band, with a responsivity of up to 0.5 A W^{-1} and a bandgap modulation of 0.2 eV for a nonplanar waveguide structure. In contrast to the graphene photodetector, which relied on a bandless structure, the dark current of the photodetector was reduced by approximately 100 times, and the noise equivalent power reached 90 pW $Hz^{-0.5}$.

In 2022, Ji et al. [37] proposed and demonstrated a scheme based on complementary polarization detection using three 90° optical mixers for self-coherent homodyne detection. The proposed complementary polarization-diversity-coherent receiver utilized remotely transmitted local-oscillator signals without optical-polarization control and had basic stability for any input polarization state of the receiver's remote local-oscillator signal.

The above literature shows that, in coherent detection, the intensity noise [16, 17, 32] caused by the local oscillator, atmospheric turbulence [10], polarization supporting role [34], spot-size and optical-axis deflection [12], coupling efficiency [20], and detector aperture size [14, 29] all have an impact on the detection sensitivity. However, through unremitting efforts, foreign scholars have reduced the impact of wavefront distortion through wavefront correction [11, 19, 21, 24, 28], reduced the impact of polarization misalignment through polarization control [33, 37], improved the coupling efficiency [20] and photodetector performance [22, 23, 25–27, 30, 31, 36], and used other methods to improve the detection sensitivity, thus approaching the limit of coherent detection.

Domestic Research Progress

In the 1980s, researchers gradually realized the important role of coherent detection in communication systems and conducted in-depth research on it. They mainly studied the influence of a series of factors on detection sensitivity, such as coupling efficiency, system thermal noise, wavefront distortion, polarization supporting role, coupling efficiency, and atmospheric turbulence.

In 1990, Da et al. [38] developed a heterodyne atmospheric laser-communication system using a CO_2 laser as the laser source and laid a solid foundation for future research on coherent optical communication systems.

In 2002, Huang et al. [39] developed a new type of photodetector to address the issue of the low reflectivity of Bragg reflectors. The detector exhibited high sensitivity and high-speed characteristics with an absorption-layer thickness of 0.2 μm. At a wavelength of 1583 nm, the peak quantum efficiency reached 80%.

In 2003, Wang et al. [40] studied the factors that affect heterodyne efficiency and found that a spatial-angle mismatch was one of the main factors. Local-oscillator

phase-plane bending and noncollimation effects also affected the heterodyne efficiency.

In 2007, Jin et al. [41] studied the statistical characteristics of Gaussian white noise and laser phase noise in homodyne-detection systems and derived the probability distribution of a two-dimensional random-variable function that took two types of noise as independent variables. Based on this, the detection error caused by the generated noise was calculated and analyzed. The results indicated that, to ensure the detection accuracy of the tracking system, it was necessary to maintain a low-laser phase-noise level and ensure high requirements for the accuracy of the phase control and tracking of the optical phase-locked loop.

In 2007, Zhao et al. [42] proposed a new receiver scheme that integrated heterodyne detection and optical-power tracking functions, established a distribution model of the detector-surface light field, and studied the heterodyne efficiency with optimal matching of the local and signal lights. Moreover, the heterodyne efficiency with incomplete matching was analyzed.

In 2008, Li et al. [43] proposed a novel adaptive-compensation method for atmospheric turbulence. The compensation effect of this method was insignificant when the relative aperture was large, for example, when $D/r_0 = 6$, the compensated intensity, only reached approximately 50% of the diffraction limit. However, when $D/r_0 = 2$, the compensation effect reached 90%. The experimental results indicated that when the relative aperture was the same, the compensation difficulty increased with an increase in turbulence intensity.

In 2008, Shu et al. [44] studied an optical microwave-subcarrier coherent-detection system and simulated and analyzed its performance. The results showed that the system was less affected by atmospheric noise and attenuation, the receiver sensitivity was high, the information-transmission quality was high, and the average bit-error rate was 4.98×10^{-9}.

In 2010, Zhao et al. [45] investigated an external-differential step demodulation method. The research results showed that the mixing efficiency was inversely proportional to the loss of the supporting-role angle. When the local-oscillator light had a thin surface bend, the mixing efficiency was directly proportional to the radius of curvature.

In 2011, Liu et al. [46] conducted in-depth research on the impact of the different amplitude distributions of the signal and local-oscillator lights on the mixing efficiency. The experimental results showed that the Airy and Gaussian distributions achieved the maximum heterodyne efficiency, whereas the heterodyne efficiency achieved by Gaussian plane waves was 82%.

In 2011, Han et al. [47] studied the compensation technology of adaptive optics for space optical communication systems based on the gamma–gamma distribution of atmospheric turbulence. The experimental results showed that adaptive-optical error-compensation technology could significantly improve the detection sensitivity, and even low-order adaptive-optical compensation achieved a good compensation effect.

In 2012, Chen et al. [48] proposed a highly sensitive detector that reduced the measurement errors caused by fluctuations in the detector light intensity and noise

carried by the signal itself. The experimental results showed that the detector had advantages, such as high sensitivity, wide operating bandwidth, and high signal-to-noise ratio, which met the research requirements of terahertz time-domain spectroscopy systems.

In 2013, Zhang et al. [49] studied the factors that affected heterodyne efficiency based on a binary phase-shift keying (BPSK) homodyne coherent-detection system. Among these, the collimation mismatch and beam waist radius had a significant impact on the heterodyne efficiency. The experimental results showed that an optimal waist beam radius could be found to maximize the mixing efficiency.

In 2013, Ma et al. [50] analyzed the relationship between the polarization state of the signal light and the mixing efficiency. The results showed that circularly polarized light had stronger stability than linearly polarized light. When both the signal light and local-oscillator light were circularly polarized signals, the mixing efficiency of the mixer was maximized.

In 2013, Ke et al. [51] selected an extruded-fiber polarization controller. The control algorithm of the system adopted a simulated-annealing algorithm, and an improved algorithm was adopted to realize automatic polarization control. The control process adopted a blind-search method. The simulation results showed that this method was suitable for polarization control in coherent optical communications.

In 2013, Xiao et al. [52] studied the coupling conditions, systems, and modes of optical fibers and analyzed the ideal coupling between spatial light and optical fibers. Three common types of fiber amplifiers were compared, and an erbium-doped fiber amplifier with the best performance was used as the front-end optical amplifier for the receiver. Subsequently, the sensitivity and signal-to-noise ratio of the coherent-detection system were analyzed.

In 2014, Hu et al. [53] developed a new type of photodetector in response to a series of requirements for coherent-light measurement systems, such as high sensitivity, wide response bandwidth, and high stability. The test results indicated that the detector could achieve remote optical-coupling transmissions and its detection sensitivity reached 0.52 μW. The coupling efficiency exceeded 38%, and the measured bandwidth reached 17.5 MHz.

In 2014, Luo et al. [54] developed an experimental platform for coherent-optical communication systems and conducted a series of experiments on the platform. The experimental results showed that when the communication distance was 600 m outdoors and the sensitivity was −17.6 dBm, the received bit-error rate was 10^{-6}. When the indoor communication distance was 6 m, the sensitivity increased by approximately 7 dB compared to that outdoors.

In 2015, Li et al. [55] studied the mixing efficiency of a coherent-detection system using a Gaussian beam model. The research results showed that the mixing efficiency reached the best value when the waist radius of the signal light was small, and the waist radius of the local-oscillator light was half of the detector radius.

In 2016, Ke et al. [56] studied the effects of local-oscillator power and detector parameters on the system signal-to-noise ratio. The experimental results showed that the characteristic parameters of the detector affected the signal-to-noise ratio and

optimal local-oscillator power. When the polarization angles of the local oscillator and signal light were larger, the optimal local-oscillator power was smaller and the signal-to-noise ratio was larger.

In 2017, China Southern Airlines et al. [57] conducted in-depth research on the impact of various factors on the mixing efficiency. They studied the error-fluctuation range of the spot-size deviation and optical-axis deflection when the mixing efficiency was greater than 10%.

In 2019, Ke et al. [58] analyzed the relationship between the intermediate-frequency signal, optical-polarization state, and mixing efficiency and developed a single-particle optimization algorithm for coherent optical communication. The experimental results showed that the mixing efficiency increased with the increase of the amplitude of the intermediate-frequency signal in the closed-loop system. When the amplitude of the intermediate-frequency signal was 0.001, the coherent-detection system could control the polarization.

In 2020, Xu et al. [59] conducted in-depth research on the design scheme and preparation process of a series of superconducting-nanowire detectors (SNDs) and developed a process for preparing superconducting-nanowire single-photon detectors with high-saturation detection efficiency. For the 1550 nm laser communication band, a six-pixel NbN superconducting-nanowire detector connected by six nanowires in series was developed. The test results confirmed that the developed six-pixel SND achieved a system-detection efficiency of up to 72.1% (dark count of 1 Hz) in the 1550 nm wavelength band, while achieving a resolution of one to six incident photons.

In 2021, Cui et al. [60] developed a new type of balanced photodetector, which consisted of a hybrid cross-resistance amplifier and a balanced photodiode chip. The experiment used a 1550 nm laser testing system. The results showed that under certain conditions, the sensitivity limit of the receiver could reach −61 dBm.

In 2021, Cheng et al. [61] analyzed the relationship between wavefront distortion and the mixing efficiency, as well as the periodic impact of wavefront distortion on the system error rate under different modulation methods. Based on the impact of wavefront distortion on the system performance, a large-amplitude wavefront-distortion residual-correction algorithm was proposed and achieved good results.

In 2022, Ke et al. [62] designed a fiber coupler with a five-degrees-of-freedom coupling device based on piezoelectric ceramics and combined it with a variable-gain random parallel gradient-descent algorithm to determine the optimal alignment posture of a spatial light single-mode fiber. The experimental results showed that the five-degrees-of-freedom fiber coupler effectively corrected different alignment errors, and the coupling efficiency between the spatial light and single-mode fiber reached 53.2% after the system was a closed loop.

In 2022, Fan et al. [63] studied the influence of the top-gate voltage on the interface coupling effect of graphene silicon photodetectors in response to the problems of a limited light-absorption capacity and the short carrier-recombination life of graphene photodetectors. The experimental results showed that when the top-gate voltage was −3 V and the light-source wavelength was 635 nm, the detector's response reached 1.1×10^4 A/W. When the wavelength of the light source

was 1550 nm, the detector's response reached 15 A/W. By increasing the top-gate voltage, the detector's response was increased five to eight times.

In 2022, Lan et al. [64] designed a highly sensitive self-driving photodetector based on a graphene/silicon nanowire-array heterostructure. This detector significantly improved the photoelectric-detection performance by increasing the effective light area and enhancing the absorption of the heterostructure. The experimental results indicated that when the incident light intensity was 90 $\mu W/cm^2$ and the wavelength was 810 nm, the response of the detector reached 0.56 $A \cdot W^{-1}$ and the response was 1.24×10^6 $V \cdot W^{-1}$.

In 2023, Xi et al. [65] adopted an easily etched and small-divergence Au/SiO_2 optical cavity, which not only improved the efficiency of nanowire light absorption but also optimized the processing accuracy of the self-aligned chip outer contour, improving the optical-coupling efficiency of the self-aligned superconducting-nanowire single-photon detector. The experiment showed that the detector could achieve a maximum detection efficiency of 82% at a temperature of 2.2 K and a laser wavelength of 1310 nm. When the wavelength was 1200–1600 nm, the system detection efficiency could reach over 65% and 40% MHz, with a 3-A count rate and a time jitter of 38 ps.

In 2023, Zhu et al. [66] proposed a transistor structure based on the high electron mobility of AlGaN/GaN and successfully developed a ZnO nanowire-sensitive gate photodetector. They were targeting the problems of the low solar-cell efficiency, low light absorption, and low photocurrent of traditional AlGaN/GaN high-electron-mobility transistor photodetectors using ZnO films. Compared with the conventional AlGaN/GaN transistor structure with high electron mobility, the peak responsivity of the new detector increased by about 2.85 times, the response time was shortened to 10 ms, and the recovery time was shortened to 250 ms, greatly improving the performance of the detector.

The above literature shows that the local oscillator [42, 56], atmospheric turbulence [54], polarization-loss supporting role [40, 45], spot-size and optical-axis deflection [57], beam waist radius [49, 55], etc. have an impact on the detection sensitivity. Therefore, increasing the wavefront correction [43, 47, 61] and polarization control [50, 51, 58], improving the coupling efficiency [62] and detector performance [39, 48, 53, 59, 60, 63–66], increasing the preamplifier [52], optimizing the detection system [44], and other methods have been proposed to improve the detection sensitivity and approach detection limits.

Factors Affecting the Coherent-Detection Limit

In a coherent-detection system, the factors affecting the detection sensitivity are as follows:

1. Thermal and shot noise
2. Wavefront distortion of the signal light
3. Polarization mismatch between the signal light and LO light

4. Spot-size deviation and optical-axis deflection
5. Coupling efficiency
6. Detector performance
7. Preoptical amplifiers

The first factor is the error caused by noise, which directly affects the detection sensitivity. The second, third, fourth, and fifth errors further affect the sensitivity of the detection system by affecting the mixing efficiency. The performance of the photodetector directly affects the sensitivity of the detection system, whereas the role of the preamplifier is to amplify a weak signal light, thereby increasing the mixing efficiency. This section elaborates on the limitations of coherent detection based on the seven aspects mentioned above.

Thermal and Shot Noise

Thermal noise is caused by the irregular movement of electrons in the conductor, and shot noise is caused by the random fluctuation of the photocurrent. This section analyzes and demonstrates the impact of thermal and shot noise on the detection sensitivity.

The total noise of the system can be described using the following equation [3]:

$$N_{\mathrm{P1}} = \frac{2e^2\eta_{\mathrm{PC}}}{h\nu} P_{\mathrm{s}} \Delta f_{\mathrm{IF}} R_{\mathrm{L}} + \frac{2e^2\eta_{\mathrm{PC}}}{h\nu} P_{\mathrm{l}} \Delta f_{\mathrm{IF}} R_{\mathrm{L}} + \frac{2\mathrm{e}^2\eta_{\mathrm{PC}}}{h\nu} P_{\mathrm{b}} \Delta f_{\mathrm{IF}} R_{\mathrm{L}}, \tag{22.3}$$

where e is the amount of electronic charge, h is Planck's constant, η_{PC} is the solar-cell efficiency, ν is the carrier frequency, P_{s} is the signal optical power, P_{l} is the local-oscillator optical power, Δf_{IF} is the effective noise bandwidth, P_{b} is the background radiation power, and R_{L} is the effective resistance of the photodetector.

The thermal noise of the system can be expressed by the following equation [3]:

$$N_{\mathrm{P2}} = 4kT_{\mathrm{b}} \Delta f_{\mathrm{IF}} \tag{22.4}$$

where N_{P2} is the thermal noise of the system, T_{b} is the operating temperature, and k is the Boltzmann constant.

When considering the thermal noise caused by the load resistances of the optical mixer and detector, the detection sensitivity is as follows:

$$\mathrm{NEP} = \frac{\left[\frac{e^2\eta_{\mathrm{PC}}}{h\nu}\left(P_{\mathrm{l}} + P_{\mathrm{b}}\right)\right] \Delta f_{\mathrm{IF}} R_{\mathrm{L}} + 2kT_{\mathrm{b}} \Delta f_{\mathrm{IF}}}{\left[\left(\frac{e\eta_{\mathrm{PC}}}{h\nu}\right)^2 P_{\mathrm{l}} R_{\mathrm{L}} - \frac{e^2\eta_{\mathrm{PC}}}{h\nu} \Delta f_{\mathrm{IF}} R_{\mathrm{L}}\right]}. \tag{22.5}$$

Research has shown that when the quantum efficiency of the detector differs, the detection sensitivity is related to the local-oscillator power, as follows: Below the quantum-noise limit, the detection sensitivity decreases as the local-oscillator power

increases. When the power of the local-oscillator remains constant, the detection sensitivity decreases as the quantum efficiency increases.

Wavefront Distortion of the Signal Light

Wavefront distortion is caused by an unequal optical path from the light source to each point on the wavefront; the magnitude of this optical-path difference quantifies the degree of wavefront distortion. Owing to the influence of atmospheric turbulence, a signal light transmitted in the atmosphere can experience unequal optical-path transmissions, resulting in the receiving end's light-wave equiphase plane becoming a randomly undulating surface with different light angles of incidence on the optical receiving surface.

If the signal and LO lights are coherently mixed without any processing, the mixing efficiency is greatly reduced, owing to the existence of a spatial mismatch, and the mixing efficiency decreases with an increase in the spatial mismatch. Figure 22.6 shows the relationship between the signal-light wavefront and LO-light wavefront on the photosensitive surface.

The output current of the photodetector is as follows [3]:

$$i_{\mathrm{IF}}(t) = G\beta E_{\mathrm{ls}} E_{\mathrm{l}} \cos[(\omega_{\mathrm{l}} - \omega_{\mathrm{s}})t + (\varphi_{\mathrm{l}} - \varphi_{\mathrm{s}})] \ \frac{\sin(\omega_{\mathrm{s}} d/2v_{\mathrm{x}})}{\omega_{\mathrm{s}} d/2v_{\mathrm{x}}}, \tag{22.6}$$

where $v_{\mathrm{x}} = c/\sin\theta$, G is the internal gain of the photodetector, β is the responsivity of the photodetector, and ω_{s}, φ_{s} and ω_{l}, φ_{s} are, respectively, the angular frequencies and phases of the signal light and the local oscillator light. E_{ls} is the signal photoelectric

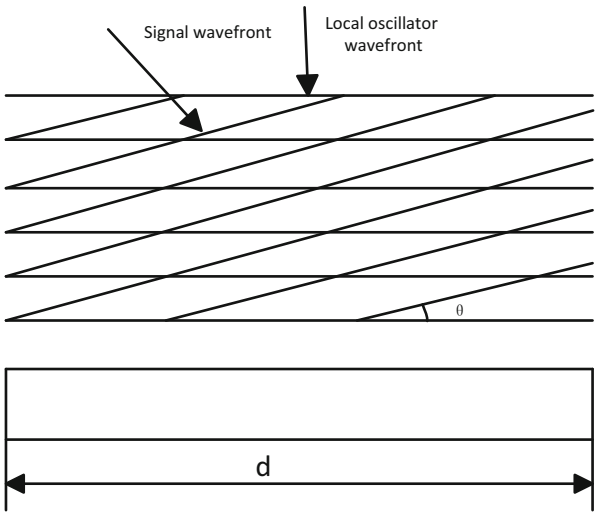


Fig. 22.6 Spatial phase relationship between the signal light and local-oscillator light [3]

field considering the wavefront distortion, d is the side length of the photosensitive surface of the detector, θ is the spatial mismatch, E_1 is the local-oscillator light field, and t is time.

In a coherent-light detection system, the mixing efficiency and difference-frequency current decrease with an increase in the spatial-adaptive angle, which directly affects the sensitivity of coherent detection. To improve the detection sensitivity, it is necessary to make the local-oscillator light and signal light as parallel and coincident on the photosensitive surface as possible. The signal light is affected by atmospheric turbulence when transmitted through the atmosphere, resulting in wavefront distortion when it reaches the receiving end. Currently, most scientists are committed to using various methods to reduce the impact of wavefront distortion on detection sensitivity. The adaptive-optics system is the most popular wavefront-correction method.

Polarization Mismatch Between the Signal Light and Local-Oscillator Light

In a general coherent optical communication system, the signal light emitted by a transmitter is generally a circularly polarized Gaussian beam. After the signal light is transmitted through the atmospheric channel, the amplitude of the optical signal received by the receiver is not a Gaussian distribution, and its polarization state is also transformed. This means that the polarization states of the signal light and the LO light are no longer consistent, leading to a reduction in the mixing efficiency and detection sensitivity.

In coherent-optical communication systems, the optical transceiver is an important component, and its function is to receive the signal light. Figure 22.7 shows the structural diagram of a typical Cassegrain-type optical antenna in an optical transceiver.

If both the intrinsic light and signal light are circularly polarized and coherently mixed after the signal light is transmitted through the atmospheric channel, the mixing efficiency can be expressed as [67]

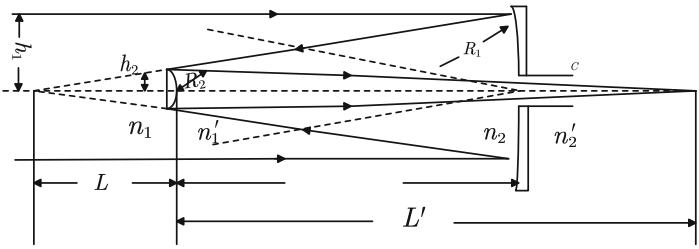


Fig. 22.7 Structural diagram of a Cassegrain-type optical antenna [67]

$$\eta = \frac{\|\cos\alpha\left|\iint\left|E_{\mathrm{s}}'\right\|E_{\mathrm{l}}\right|\sim \mathrm{d}s|^2}{\iint|E_{\mathrm{l}}|^2 \sim \mathrm{d}s \iint\left|E_{\mathrm{s}}'\right|^2 \sim \mathrm{d}s}, \tag{22.7}$$

where α is the polarization-mismatch angle between the signal light and local-oscillator light, E_{s}' is the signal light field, E_{l} is the local-oscillator light field, s is the photosensitive area of the photoelectric receiver, and η is the mixing efficiency.

According to Eq. (22.7), the mixing efficiency is proportional to cos α. When the value is 0–180°, it is inversely proportional to η, and the mixing efficiency η is inversely proportional to $|\cos\alpha|^2$.

In summary, it can be seen that the polarization mismatch between the local oscillator and signal light has a significant impact on the mixing efficiency. In practical engineering, polarization-control compensation devices should be used to maintain the polarization state of the signal light and local-oscillator light as consistent as possible, which is very important for improving the detection sensitivity.

Spot-Size Deviation and Optical-Axis Deflection

The deviation in spot size is caused by errors in the optical-component manufacturing process, leading to the divergence or contraction of the incident beam. Consequently, it does not match the size of the spot generated by the local-oscillator light. To obtain high-quality beams, spatial light mixers often place the receiving target surface at the waist of the incident beam; therefore, the spot size is often described by the waist radii of the intrinsic and signal beams, and the mixing efficiency formula is obtained [68]:

$$\eta_{\mathrm{oularrur}} = \frac{\left[\int_0^n \exp\left(-\frac{r^2}{r_x}\right)\exp\left(-\frac{r^2}{r_{lo}^2}\right) r\,\mathrm{d}r\right]^2}{\int_0^{r_0}\left[\exp\left(-\frac{r^2}{r_{lo}}\right)\right]^2 r\,\mathrm{d}r \int_0^{\infty}\left[\exp\left(-\frac{r^2}{r_x}\right)\right]^2 r\,\mathrm{d}r}, \tag{22.8}$$

where r is the radius of the receiving target surface, r_{x} and r_{lo} are the waist radii of the signal light and local-oscillator light, respectively, and exp (·) is an exponential function with e as the base.

Processing errors in the optical components may cause the incident beam to diverge and compress, resulting in different spot sizes for the signal light and local light at the receiving target surface. Using the spot size of the local oscillator as a reference, the deviation in the spot size can be calculated as $\Delta r = r_{\mathrm{x}} - r_{\mathrm{lo}}$. Substituting the deviation into optical mixing-efficiency formula (22.8) yields a new optical mixing-efficiency formula with the spot-size deviation [68]:

$$\eta_{\text{colarert}} = \frac{\left[\int_0^{r_0} \exp\left(-\frac{r^2}{(r_{\text{lo}}+\Delta r)^2}\right) \exp\left(-\frac{r^2}{r_{\text{lo}}^2}\right) r\,\mathrm{d}r\right]^2}{\int_0^n \left[\exp\left(-\frac{r^2}{r_{\text{lo}}^2}\right)\right]^2 r\,\mathrm{d}r \int_0^\infty \left[\exp\left(-\frac{r^2}{(r_{\text{lo}}+\Delta r)^2}\right)\right]^2 r\,\mathrm{d}r}, \tag{22.9}$$

where Δr is the spot-size deviation.

In practical applications, both displacements and size mismatches of the light spot exist on the receiving target surface because the processing and adjustment errors of the optical components inside the spatial light mixer cannot be eliminated. Therefore, both of these factors must be considered simultaneously. Therefore, the mixing-efficiency expression of the mixer when both factors exist simultaneously is obtained as follows [69]:

$$\eta_{\text{ls}} = \frac{\left[\int_0^n \exp\left(-r^2/r_{\text{lo}}^2\right) \cdot \exp\left(-r^2/(r_{\text{lo}}+\Delta r)^2\right) \cdot J_0\left(C_{\text{p}} k r \delta\right) r\,\mathrm{d}r\right]^2}{\left[\int_0^{r_0} \left[\exp\left(-r^2/r_{\text{lo}}^2\right)\right]^2 r\,\mathrm{d}r\right]\left[\int_0^\infty \left[\exp\left(-r^2/(r_{\text{lo}}+\Delta\omega)^2\right)\right]^2 r\,\mathrm{d}r\right]}, \tag{22.10}$$

where $J_0(\cdot)$ is the zero-order Bessel function, r is the radius of the receiving target surface, δ is the deflection angle of the optical axis, k is Boltzmann's constant, and C_{p} is a constant.

From the above equation, it can be seen that deviations in the spot size and optical-axis deflection can affect the mixing efficiency of the mixer to some extent. In practical coherent-light detection, methods have often been adopted to suppress this error. Commonly used methods include improving the level of hardware processing or determining the relationship between the spot-size deviation, optical-axis deflection, and mixing efficiency to maximize the mixing efficiency.

Coupling Efficiency

The ability of optical fibers to couple signal light affects the sensitivity of coherent detection systems to a certain extent. This ability to couple optical fibers is called the coupling efficiency of fiber couplers. Figure 22.8 shows a schematic of the fiber-optic coupler. When all conditions are ideal, the light rays converge at the focal point of the lens after passing through the lens and then advance along the direction of the axis in the fiber until they reach the next module.

The coupling efficiency is expressed as follows:

$$\eta_{\text{Coupling}} = P/P_0, \tag{22.11}$$

where P represents the optical power entering the fiber and P_0 represents the total incident power.

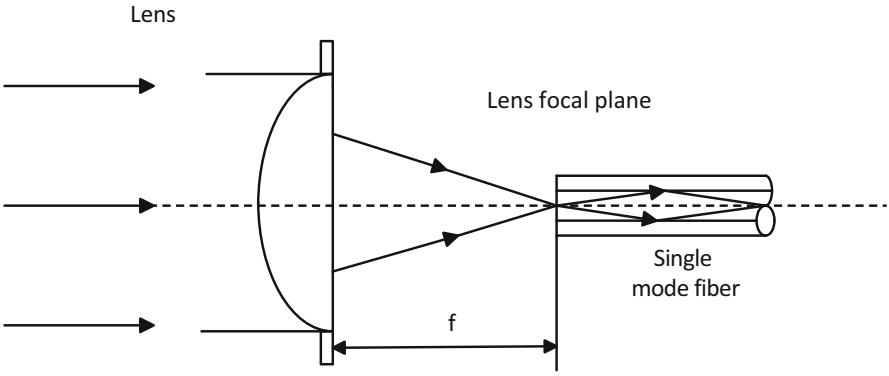


Fig. 22.8 Schematic diagram of a fiber-optic coupling

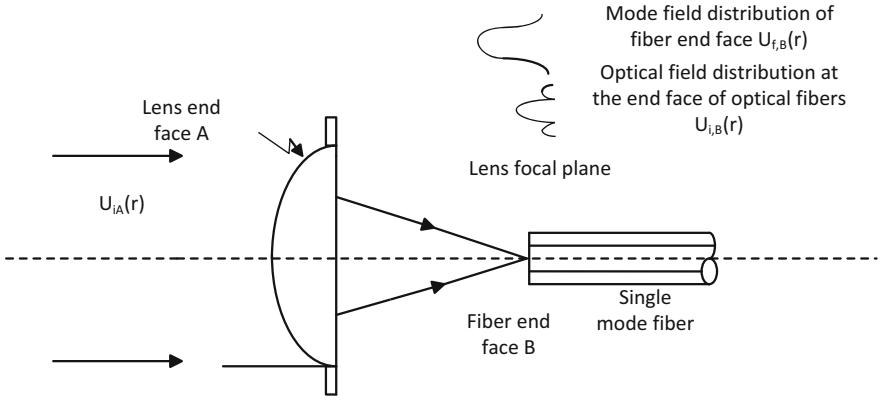


Fig. 22.9 Schematic diagram of a fiber-coupling mode-field analysis

As shown in Fig. 22.9, the coupling efficiency of the fiber coupler η_{Coupling} can be expressed as follows [70], when analyzed using the mode-field matching method:

$$\eta_{\text{Coupling}} = \frac{\left|\iint U_{\text{i}}^{*}(r)U_{\text{f}}(r)r\,\mathrm{d}r\,\mathrm{d}\gamma\right|^{2}}{\iint U_{\text{i}}(r)U_{\text{i}}^{*}(r)r\,\mathrm{d}r\,\mathrm{d}\gamma \times \iint U_{\text{f}}(r)U_{\text{f}}^{*}(r)r\,\mathrm{d}r\,\mathrm{d}\gamma}, \tag{22.12}$$

where $U_{\text{i}}(r)$ is the amplitude distribution of the signal light, $U_{\text{i}}^{*}(r)$ is the conjugate function of $U_{\text{i}}(r)$, $U_{\text{f}}(r)$ is the distribution of the electromagnetic field of a single-mode fiber on the lens end face, γ is the conjugate function of $U_{\text{f}}(r)$, r is the radius of the lens end face, and η_{Coupling} is the coupling efficiency.

In summary, it can be seen that the coupling efficiency of fiber-optic couplers is influenced by a series of factors, such as lens parameters, reception angle, signal-light amplitude, etc. In future research, these factors should be improved to enhance

the coupling efficiency of the couplers and increase the sensitivity of coherent detection.

Detector Performance

A photodetector converts optical signals into electrical signals. In the study of coherent-detection limits, the detection sensitivity largely depends on the performance of the detector. The detectors commonly used at present include photomultiplier tubes, single-photon detectors, and superconducting detectors. Figure 22.10 shows a schematic diagram of the photoelectric effect.

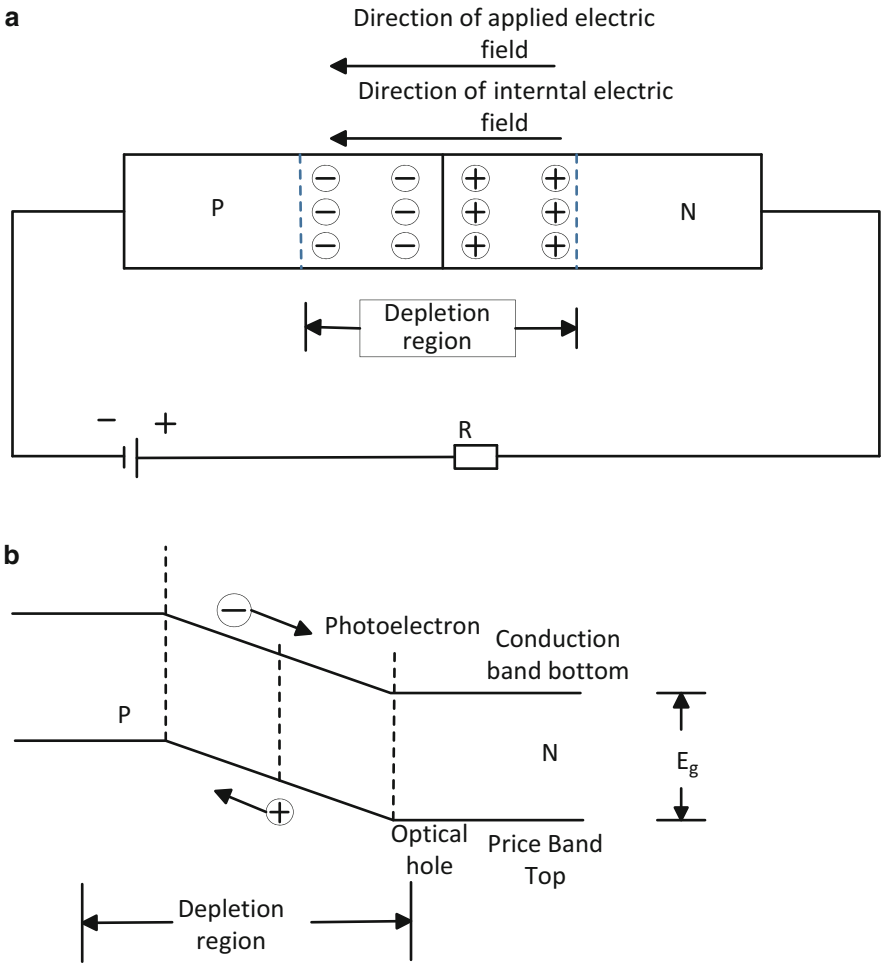


Fig. 22.10 Schematic diagram of the photoelectric effect: (**a**) Structure diagram of the depletion region of the P-N junction; (**b**) energy-band diagram around the P-N depletion region

Figure 22.10a shows the change in the P-N junction depletion region when the externally applied electric field is in the same direction as the internal electric field. Figure 22.10b shows the energy-band diagram near the P-N junction depletion region. It can be seen that the electrons move in the energy band, the potential of the location where the electrons gather is low, and the energy band is high.

With the development of science and technology, new detectors with higher sensitivities and faster responses, such as nanowire and graphene detectors, have emerged in recent years. Nanowire detectors use semiconductor-nanowire materials for material detection and sensing. A charge transfer can occur when the target material reacts with the surface of the nanowire, thereby changing its electrical properties and facilitating detection.

Graphene photodetectors utilize the electronic and optical properties of graphene for photoelectric conversion. The electronic properties of graphene enable it to have a high-speed electron-transmission capability, high sensitivity, high transmittance, and low reflectivity. Improvements in various aspects of the photodetector performance are of great significance for improving the sensitivity of coherent detection.

Preoptical Amplifier

A preamplifier is commonly used in optical communications and sensor systems. Its primary function is to amplify a weak optical signal at the input end of an optical fiber or detector with a sufficiently high signal-to-noise ratio. The device does not make any internal photoelectric or electro-optical conversions.

In coherent-detection systems, optical amplifiers are an essential device that plays a significant role in improving the detection sensitivity and increasing detection limits. There are two types of commonly used optical amplifiers: semiconductor and fiber. This section introduces an erbium-doped fiber amplifier (EDFA) with excellent performance.

As shown in Fig. 22.11, the structural diagram of the EDFA consists of two parts: an optical path and an auxiliary circuit. The optical path is composed of an optical isolator, pump light source, optical coupler, erbium-doped fiber, and optical filter. The auxiliary circuit is composed of a power supply, warning and protection circuit, and automatic microprocessor control. The auxiliary circuit ensures the normal operation of the optical amplifier. The pump source provides sufficient energy for the erbium-doped fiber to invert the particle beam.

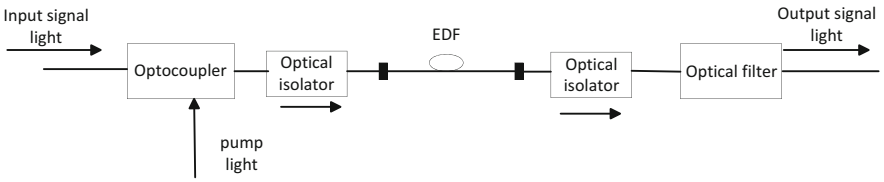


Fig. 22.11 Structural block diagram of an EDFA

The core device of the optical amplifier is an erbium-doped fiber, and its main material is quartz fiber. In practice, to improve the performance of an optical amplifier, the rare-earth element Er^{+3} is often doped inside the quartz.

In digital optical-communication systems, regardless of the circuit thermal noise and extinction ratio, the detection sensitivity can be expressed as follows:

$$P_s = 2h\nu n_{sp} B_e \left(Q^2 + Q \cdot \sqrt{0.5\mu B_0 / B_e - 0.5} \right), \tag{22.13}$$

where h is Planck's constant, η_{sp} represents the spontaneous emission factor, B_0 represents the bandwidth of the optical filter, B_e is the bandwidth of the electric filter, and ν is the carrier frequency. The value of μ is 1 or 2 and depends on whether a polarization filter is used.

When the bit error rate is 10^{-9}, $Q = 6$; when the bit-error rate is 10, $Q = 5.2$. If the receiving end adopts error-correction coding and the preamplifier uses a silver-doped fiber amplifier, the sensitivity of coherent detection can reach -50 dBm.

Development Trends

Coherent-detection technology is used to measure coherent properties and is one of the most important technologies in the field of optics. With the continuous progress in science and technology, coherent-detection technology is constantly developing. This section provides a detailed explanation of the development trends in coherent-detection technology.

High-Detection Sensitivity and Resolution

The highest sensitivity of coherent detection is achieved by the single-photon counting of photon counters or single-photon detectors. However, the resolution of this method is extremely low. Therefore, higher-sensitivity methods are required to improve the resolution of coherent detection. The emergence of new technologies has improved the sensitivity and resolution of coherent detection. The main approach is to use higher detector sensitivity, more efficient photon splitters, and better coherent interferometers.

Wider Applications

Coherent-detection technology has wide applications in various fields, such as optical communication, military, optical imaging, and microscopy. In the future, with the development of coherent-detection technology, it will gradually be applied to other fields, such as quantum computing, high-speed communication, and biomedicine. The demand for efficient and accurate coherent-detection technology in

these fields is increasing; therefore, the application of coherent detection technology will become increasingly widespread.

Simpler and More Efficient Detection Methods

With the continuous development of science and technology, methods for coherent-detection technologies are constantly being developed. In the future, this equipment will become more efficient and precise. The emergence of new coherent-detection technologies has made coherent detection more convenient. For example, the automated control technology of optical interferometers can significantly improve the efficiency of coherent detection, reducing its implementation difficulty.

Further Research

Coherent-detection technology is an active research field, and further research may be conducted in the future to expand its application fields and improve the performance of detection technology. With the rapid development of communication technologies, research on coherent-detection technologies has received increasing attention and will likely continue to receive the attention of scientists and researchers.

In summary, coherent-detection technologies will continue to develop and play important roles in many fields. The development trend of future coherent-detection technologies will include higher detection sensitivities, wider application fields, more efficient and simplified detection methods, and deeper research. With the advancement of these technologies, coherent-detection technologies will become increasingly indispensable in modern optical fields.

Future Prospects

This chapter summarized the research progress in coherent detection, both domestically and internationally. Coherent detection can be divided into heterodyne and homodyne detection, according to the detection method. Then, the factors affecting the detection sensitivity were analyzed:

1. Thermal noise and shot noise generated by the system
2. Wavefront distortion of the signal light
3. Polarization mismatch between the signal light and LO light
4. Spot-size deviation and optical-axis deflection
5. Optical-fiber coupling efficiency
6. Performance of the photodetector

Suppression schemes have been developed to address these effects and achieve the highest detection sensitivity. Although the current coherent-detection limit technology is relatively mature, the following issues are still worth studying and discussing:

1. Use better laser sources. By using a laser source with a higher power, narrower linewidth, and higher spot-quality factor, the coherence of the laser beam can be improved, thereby improving the detector sensitivity.
2. Use higher detection efficiency. Maximizing the detector response efficiency can improve the sensitivity and accuracy. Common methods include increasing the detector area, increasing the detector gain, and improving the detector response time.
3. Reduce noise. The use of low-noise preamplifiers and signal processors to reduce noise can improve the sensitivity and accuracy of detectors.
4. Optimize the detector structure and materials. Changing the process, structure, and materials of a detector can improve its sensitivity and accuracy. For example, new semiconductor materials can be used to increase the doping concentration and change the detector structure.
5. Use adaptive-optics technology. By using adaptive-optics technology, the phase perturbation of coherent beams can be eliminated, thereby improving the sensitivity and accuracy of coherent detectors.

References

1. Gudmundsen R A. Photoelectric mixing of incoherent light waves [J]. Physical Review, 1955, 99(6): 1691–1700.
2. Forrester A T, Gudmundsen R A, and Johson P O. Photoelectric mixing of inherent light [J]. Physical Review, 1955, 99(6): 1691.
3. Ke Xizheng, Wu Jiali. Principles and Applications of Wireless Optical Coherent Communication [M] Beijing: Science Press, 2019.
4. Zhu Ling, Feng Ying. Research progress and trends in satellite coherent optical communication [J]. Progress in Laser and Optoelectronics, 2007, 44 (6): 49–53.
5. Yao Zhoushi, Hu Yu. Development History and Trends of Intersatellite Coherent Optical Communication Technology [J]. Optical Communication Technology, 2005, 29 (08): 44–46.
6. Sodnik Z F B L H. Optical Intersatellite Communication [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2010, 16(5):1051–1057.
7. Ke Xizheng, Deng Lijun. Wireless Optical Communication [M]. Beijing: Science Press, 2016, 7–17.
8. Liu C, Chen S Q, Li X Y, et al. Performance evaluation of adaptive optics for atmospheric coherent laser communications [J]. Optics Express, 2014, 22(13):15554–15563.
9. Barry J R, Lee EA. Performance of coherent optical receivers [J]. Proceedings of the IEEE, 1990, 78(8):1369–1394.
10. Teich, M. C. OPTIMUM HETERODYNE DETECTION AT 10.6 μm IN PHOTOCONDUCTIVE Ge:Cu [J]. Applied Physics Letters, 1966, 9(10):357–360.
11. Popoola W o, Ghassemlooy Z. Performance of subcarrier modulated freespace optical comunication link in negative exponential atmospheric turbulence environment [J]. International Journal of Autonomous and Adaptive communications Systems, 2008, 1(3): 342–355

12. Cohen S C. Heterodyne detection: phase front alignment, beam spot size, and detector uniformity [J]. Appl Opt, 1975, 14(8):1953–1959.
13. Abbas G L, Chan V, Yee T K. Local-oscillator excess-noise suppression for homodyne and heterodyne detection [J]. Optics Letters, 1983, 8(8):419–421.
14. Tanaka K, Ohta N. Effects of tilt and offset of signal field on heterodyne efficiency [J]. Applid Optics, 1987, 26(4): 627–632.
15. Calvani R, Caponi R. Polarisation phase-shift keying: a coherent transmission technique with differential heterodyne detection [J]. Electronics Letters, 1988, 24(10):642–643.
16. Tanaka, Taguchi. Heterodyne efficiency for a partially coherent optical signal. [J]. Applied optics, 1992, 46(56):789–799.
17. Bar-David I, Salz J. On dual optical detection: homodyne and transmitted-reference heterodyne reception [J]. IEEE Transactions on Communications, 2002, 36(12):1309–1315..
18. Wang L, Choa F S. Counter receiving Heterodyne Detection with an Integrated Coherent Transceiver and Its Applications in Bandwidth-On-Demand Access Networks [J]. Journal of Lightwave Technology, 1999, 17(10):P.1724–1731.
19. Sato M, Onodera K, Akiba M, et al. Phase-drift suppression in heterodyne detection and its application to optical coherent tomography [J]. Filtration Industry Analyst, 2000, 41(15): 388–391.
20. Gol'tsman, G. N, Okunev, et al. Picosecond superconducting single-photon optical detector. [J]. Applied Physics Letters, 2001, 25(8), 115:125.
21. Ricklin J C. Free-space laser communication using a partially coherent source beam. [D]. The Johns Hopkins University. 2002.
22. Sobolewski, Roman, Verevkin, et al. Ultrafast Superconducting Single-Photon Optical Detectors and Their Applications. [J]. IEEE Transactions on Applied Superconductivity, 2003, 13(2): 1151–1157.
23. Fernandez M M, Vilnrotter V A. Coherent optical receiver for PPM signals received through atmospheric turbulence: Performance analysis and preliminary experimental results [J]. Optical Engineering, 2004, 56(62), 5338–5345.
24. Yang Y, Lu Y, Chen Y, et al. A radial-shearing interference system of testing laser-pulse wavefront distortion and the original wavefront reconstructing [J]. Optical Engineering, 2005, 5638(1):200–204.
25. Dauler E A, Robinson B S, Kerman A J, et al. Multi-Element Superconducting Nanowire Single-Photon Detector [J]. IEEE Transactions on Applied Superconductivity, 2007, 17(2): 279–284.
26. Hyundai, Park, Alexander, et al. A hybrid AlGaInAs-silicon evanescent waveguide photodetector. [J]. Optics Express, 2007, 15(10): 6044–6052.
27. Xu Q, Arvizu Mondragón, A, Gallion P, et al. Homodyne In-Phase and Quadrature Detection of Weak Coherent States With Carrier Phase Tracking [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2009, 15(6):1581–1590..
28. Gregory M, Hemmati H, Heine F, et al. Inter-satellite and satellite-ground laser communication links based on homodyne BPSK [J]. Proc Spie, 2010, 65(68), 7587–75870.
29. Salem M, Rolland J P. Heterodyne efficiency of a detection system for partially coherent beams [J]. Journal of the Optical Society of America A, 2010, 27(5): 111–119.
30. Gan X, Shiue R J, Gao Y, et al. Chip-integrated ultrafast graphene photodetector with high responsivity [J]. Nature Photonics, 2013, 7(11):883–887.
31. Boroson D M, Robinson B S. The lunar laser communication demonstration: NASA's first step toward very high data rate support of science and exploration missions [J]. Space Science Re-views, 2014, 185(4): 115–128.
32. Chengqiang L, Tingfeng W, Heyong Z, et al. The performance of heterodyne detection system for partially coherent beams in turbulent atmosphere [J]. Optics Communications, 2015, 356(56):620–627.

33. Wang Y, Li c Q, Wang T F, et al. The effects of polarization changes of stochastic electromagnetic beams on heterodyne detection in turbulence [J]. Laser Physics Letters, 2016, 13(11): 116–126.
34. Yang Y F, Yan C X, H C H, et al. Modified heterodyne efficiency for coherent laser communication in the presence of polarization aberrations [J]. Optics Express, 2017, 25(7): 7567–7591.
35. Chen M, Liu C, Rui D, et al. Highly sensitive fiber coupling for free space optical communications based on an adaptive coherent fiber coupler [J]. Optics Communications, 2019, 430(87): 223–226.
36. Maiti R, Patil C, Saadi M, et al. Strain-engineered high-responsivity MoTe2 photodetector for silicon photonic integrated circuits [J]. Nature Photonics, 2020, 14(9):578–586.
37. H. Ji. Complementary Polarization-Diversity Coherent Receiver for Self-Coherent Homodyne Detection With Rapid Polarization Tracking [J]. Lightwave Technology, 2022, 64(8): 2773–2779.
38. Nanjing Da, Pi Mingjia, Fan Liming. CO_ Study on the Sensitivity of 2 Laser Heterodyne Detection Systems [J]. Journal of Optics, 1990, 10 (08): 714–720.
39. Huang Hui, Wang Qi, Lei Lei,. Long wavelength and high sensitivity InP/InGaAs resonant cavity photodetectors [J]. Optoelectron Laser, 2002, 13(03): 221–224.
40. Wang Qi, Wang Chunhui, Shang Tieliang. Heterodyne efficiency of coherent detection of Gaussian local oscillator light and Airy spot signal light [J] China Laser, 2003, 30 (89): 183–186.
41. Jin Tao, Gu Lei. Research on detection error of coherent detection and orientation system [J]. Acta Photonica Sinica, 2007, 36 (02): 340–343.
42. Zhao Changzheng. Preliminary Scheme and Heterodyne Efficiency Study of Spatial Coherent Optical Communication [D] Qinhuangdao: Yanshan University, 2007
43. Li Huan, Zhang Hongtao, Yin Fuchang. Adaptive compensation method for atmospheric turbulence in space laser communication systems [J]. Journal of Changchun University of Technology (Natural Science Edition), 2008, 98 (02): 1–3.
44. Cheng Shu, Ji Hang, Ma Yong. Design and simulation study of optical microwave subcarrier multiplexing coherent detection system [J]. Ship Electronics Engineering, 2008, 56 (02): 61–165.
45. Zhao Chunying. Research on Heterodyne Asynchronous Demodulation Technology for Coherent Optical Communication [D]. Changchun: Changchun University of Technology, 2010.
46. Liu Hongzhan, Ji Yuefeng, Xu Nan. The influence of amplitude distribution of signal light and local oscillator light on the mixing efficiency of inter satellite coherent optical communication systems [J]. Journal of Optics, 2011, 31 (10): 63–68.
47. Han Liqiang, Yu Qi, etc. Study on the performance of free space optical communication system based on adaptive optics compensation [J]. Applied Optics, 2010, 31 (2): 301–303.
48. Chen Longchao, Fan Wenhui Research on High Sensitivity and Low Noise Terahertz Electro-optic Detectors [J] Journal of Electronics, 2012, 40 (9): 1705–1709.
49. Zhang Tong, Tong Shoufeng. The effect of collimation mismatch on the mixing efficiency of spatial coherent laser communication [J]. Journal of Changchun University of Technology (Natural Science Edition), 2013, 36 (Z1): 13–15.
50. Ma Tingting, Tong Shoufeng, Nan Hang, et al. The influence of signal polarization characteristics on the mixing efficiency of spatial coherent detection [J]. Progress in Laser and Optoelectronics, 2017, 54 (2): 20–64.
51. Li Tie, Ke Xizheng, Chen Juan. Polarization control in coherent light detection systems [J]. Infrared and Laser Engineering, 2012, 41 (11): 3069–3074.
52. Xiao Xiang Space Light Receiving Technology Based on Pre amplifier [D]. Changchun University of Technology, 2013.
53. Hu Pengcheng, Yang Hongxing, Mei Jianting. Research on High Sensitivity Detectors for Long Range High Speed Laser Heterodyne Signals [J]. Optoelectronics. Laser, 2014, 25 (05): 925–931.

54. Luo Binbin. Space Wireless Digital Coherent Optical Communication [D]. Beijing: Beijing University of Posts and Telecommunications, 2014.
55. Li Xiangyang, Ma Zongfeng, Shi Dele. Influence of Gaussian beam field distribution on coherent detection mixing efficiency [J]. Infrared and Laser Engineering, 2015, 44 (02): 539–543.
56. Kong Yingxiu, Ke Xizheng, Yang Yuan. The influence of local oscillator power on signal-to-noise ratio in spatial coherent optical communication [J]. Infrared and Laser Engineering, 2016, 45 (02): 242–247.
57. Southern Airlines, Zhang Peng, Tong Shoufeng. Analysis of the Influence of Spot Size Deviation and Optical Axis Deflection on the Mixing Efficiency of Space Optical Mixers [J]. Infrared and Laser Engineering, 2017, 46 (04): 205–212.
58. Kong Yingxiu, Ke Xizheng, Yang Yuan. The effect of laser linewidth on the performance of spatial coherent optical communication systems [J]. Journal of Instrumentation, 2017, 38 (07): 1668–1674.
59. Tao Xu High efficiency and high-speed characteristics of superconducting nanowire single photon detectors [D]. Nanjing University, 2020.
60. Cui Dajian, Zhou Lang, Xi Shuiqing, Ren Li, Lan Feng, Guo Jingzhong. Research on Spaceborne High Sensitivity Balanced Photodetectors [J]. Semiconductor Optoelectronics, 2020, 41 (04): 480–484.
61. Cheng Shuang Research on Correction Technology for Large Amplitude Distorted Wavefront in Free Space Coherent Optical Communication [D]. Xi'an: Xi'an University of Technology, 2021.
62. Wu Jiali, Ke Xizheng, Yang Shangjun. Calibration of alignment errors in spatial light single mode fiber coupling using multidimensional couplers [J]. Journal of Optics, 2022, 42 (07): 40–50.
63. Fan Anqi Research on graphene composite structure photodetectors [D]. Chengdu: University of Electronic Science and Technology, 2022.
64. Lan Zhenli, He Feng, Song Yiji, Ding Ding, Zhou Guofang. High performance graphene/silicon nanowire array heterojunction photodetector [J]. Laser and Infrared, 2022, 52 (11): 1671–1677.
65. Xi Lingling, Yang Xiaoyan, Zhang Tianzhu. High comprehensive performance superconducting nanowire single photon detector [J]. Journal of Physics, 2023: 1–18 [2023-05-09]
66. Zhu Yanxu, Li Jianwei, Li Qixuan,. ZnO nanowire photosensitive gate photodetector based on AlGaN/GaN HEMT structure [J]. Journal of Beijing University of Technology, 2023, 49 (02): 188–196.
67. Zhu Weixia, Wan Jiayou. The effect of polarization mismatch on the performance of spatial coherent optical communication systems [J]. Semiconductor Optoelectronics, 2018, 39 (03): 435–439.
68. Nan Hang, Zhang Peng, Tong Shoufeng. Analysis of the Influence of Spot Size Deviation and Optical Axis Deflection on the Mixing Efficiency of Space Optical Mixers [J]. Infrared and Laser Engineering, 2017, 46 (04): 205–212.
69. Ke Xizheng, Wu Gali. Principles and applications of wireless optical coherent communication [M]. Beijing: Science Press, 2019, 37–47.
70. W. Zhu, J.Y. Wan. Impact of polarisation mismatch on the performance of spatially coherent optical communication systems [J]. Semiconductor Optoelectronics, 2018, 39 (03): 435–439.

Adaptive Optics Technology

23

Contents

Introduction 830
Foreign Research Progress 830
Domestic Research Progress 834
Adaptive Optics in the Optical Wireless Communication Field 838
Adaptive Optical System with Wavefront Measurement 840
Adaptive Optical System Without Wavefront Measurement 846
Wavefront Correction Using Liquid Crystal Spatial Light Modulators 856
Combined Wavefront Correction Using Tilt and Deformable Mirrors 857
Spatial Optical Fiber-Coupled Adaptive Optical Wavefront Correction 859
Summary 866
References 866

Abstract

Optical wireless coherent communication is a wireless transmission optical communication technology that uses the interference principle of light to coherent demodulate the received signal using a local oscillator light source at the receiver and compared with intensity modulation/direct detection. It has a detection sensitivity of about 20–23 db and is more suitable for long distance laser communication transmission. The existence of atmospheric turbulence limits the development of optical wireless coherent communication technology. Adaptive optics technology can effectively suppress the wavefront distortion caused by atmospheric turbulence by means of the phase conjugate compensation principle. In this chapter, the adaptive optics technologies are applied to the field of optical wireless coherent communication by taking into account the above research background.

Keywords

Optical communication · Coherent optical communication · Automatic frequency control · Decision control circuit · Narrow linewidth laser

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_23

Introduction

Adaptive optics is a comprehensive science integrating optoelectromechanics, which can suppress the disturbance of laser signals in turbulent atmospheric transmissions [1]. It is a theory, system, technology, and engineering field that studies the real time improvement of optical wavefront quality.

Adaptive optics covers the study of optics, communication, control, computers, machinery, and other disciplines. It intends to correct the random wavefront distortions caused by changes in the external environment during beam propagation in real time and automatically improves the beam quality. Since Babcock proposed the concept of adaptive optics in 1953 [2], this technology has been widely used in astronomical observations [3], human eye imaging and vision research [4], inertial confinement fusion [5], and other fields.

Foreign Research Progress

Optical wireless communication is affected by atmospheric turbulence. As an optical signal propagates, its wavefront is affected by turbulence and generates random fluctuations, which cause beam expansion, angle of arrival fluctuations, beam bending, and beam drift. These seriously interfere with the quality of the received optical signal, resulting in increased communication errors and reduced communication stability.

In addition, with an increase in the communication distance, the laser output power is gradually increased, and the laser cavity mirror undergoes a thermal deformation, which causes a phase distortion of the laser wavefront, leading to a reduction in the beam quality, and thus, the communication quality. Therefore, methods are needed to suppress the influence of atmospheric turbulence and improve the quality of laser communication. Applying adaptive optics technology to optical wireless communication systems [6] has great potential for suppressing the influence of atmospheric turbulence.

In 1972, the Fried's Optical Sciences Company of the United States, among others, introduced adaptive optics technology into optical wireless communication systems. The adaptive optics unit corrected the wavefront distortion caused by atmospheric turbulence in real time [7].

In 1996, Tyson of the University of North Carolina in the United States analyzed the relationship between light intensity fluctuation and the bit error rate in the intensity modulation/direct detection (IM/DD) mode [8, 9]. When the atmospheric turbulence intensity $D/r_0 = 0.8$ (r_0 is the atmospheric coherence length), the system bit-error rate was 2.3×10^{-3} before the wavefront phase correction and 5.5×10^{-4} after correction [10]. The wavefront distortion caused by atmospheric turbulence increased the bit-error rate by one order of magnitude.

In 2002, Wilks and others at the Lawrence Livermore National Laboratory of the United States conducted numerical simulations of optical communication with a transmission rate of 10 Gbit/s and a transmission distance of 33.2 km. After the

wavefront distortion was corrected, the coupling efficiency from space light to optical fiber was increased from 20% to 55% [11]. The experimental platform used a 400 unit microelectromechanical (MEMS) deformable mirror and a wavefront detector with a 1-kHz sampling frequency to conduct ground–space laser communication experiments on horizontal and slanted links. After correction, the root mean square (RMS) residual of the wavefront was 0.75λ [12].

In 2004, Thomas et al. of the University of Maryland in the United States used tilt mirrors to correct tilt aberrations, and 132 units of MEMS deformable mirrors to correct other aberrations. After correcting the higher order aberrations, the received signal strength was twice as high as when only tilt aberrations were corrected [13, 14].

In 1990, Wright et al. of the Jet Propulsion Laboratory (JPL) built an optical wireless communication experimental platform based on adaptive optics using a 1064 nm laser. When the system transmission rate was 100 Mbit/s, the gain of the signal received by the receiving end after wavefront correction was approximately 6 dB [15].

In 2006, Hemmati et al. of the JPL used deformable mirrors to actively compensate for received wavefront distortion. The receiving aperture was 0.3 m, and the incident beam wavelength was 632 nm. The root mean square error of the wavefront distortion was 1.4λ, which decreased to 0.05 λ. The Strehl ratio (SR) increased from 0.08% to 89% [16].

In 2010, Wilson et al. of the JPL conducted a demonstration experiment of Mars deep-space laser communication with a 1064 nm communication wavelength and a pulse-position modulation communication mode. The adaptive optical system in the communication system used a 635 nm band as a beacon for wavefront detection and correction, with a communication rate of 100 Mbit/s. After correction, the error rate was reduced by 3 dB [17, 18].

In 2013, Stewart et al. of the National Aeronautics and Space Administration (NASA) of the United States experimentally verified laser communication between the earth and the moon. The system used a MEMS deformable mirror to obtain a correction result with an SR of 0.6 [19].

In 2015, Wright et al. of the JPL tested an adaptive optics system at a ground observation station with an optical communication telescope. The system received a 50 Mbit/s downlink video transmission from the International Space Station, and the SR of the beam received by the satellite increased from 0.03 to 0.66 after correction [20].

In 2008, the Johns Hopkins University Applied Physics Laboratory and AOptix Technologies demonstrated a 147 km long range bidirectional radio frequency (RF) and optical wireless communication experiment using intensity modulation at transmission rates of up to 10 Gbit/s. The system included a 35 driver deformable mirror to provide a bandwidth greater than 1 kHz. The coupling efficiency between the signal light and the single-mode fiber (SMF) was improved by adding transmitting terminal pre corrections and receiving terminal corrections. It provided 30 closed loop corrections for Zernike aberrations, in addition to the tilt and focus terms [21].

In May 2010, Heine et al. from the European Space Agency installed and commissioned an adaptive optical system at a ground station for satellite based laser communication. The working wavelength was 1064 nm, the sensor was an 88 subaperture Hartmann sensor, and a 12 × 12 unit MEMS deforming mirror control unit was used. After acquisition by a high speed infrared camera, the sampling frequency of the maximum subaperture image reached 20 kHz [22–24].

In 2011, Gregory et al. of the European Space Agency used a 1-m aperture optical system at a ground station in Spain to conduct an adaptive optical correction experiment for satellite based laser communication. The communication wavelength of the adaptive optical system was 1064 nm, and the SR was increased to more than 0.4 after correction [25].

In 1995, Arimoto et al. from Japan's Ministry of Posts and Telecommunications conducted a study on laser communication between the ground and the International Space Station (ISS). After compensating the wavefront distortion of the signal light with the ground-end adaptive system, the coupling power of the downlink increased from −58.9 dBm to −51.9 dBm [26].

In 1999, Kudielka et al. from Japan's Ministry of Posts and Telecommunications conducted laser communication experiments between the ISS and the ground. After correcting low order aberrations, such as tilt, the gain of the optical power was increased by approximately 2.3 dB [27]. In 2016, Petit et al. used ground optical receivers, small optical communication terminals, and satellite laser communication terminals to conduct laser communication experiments with adaptive optical correction; the error suppression bandwidth was less than 60 Hz [28].

In 2014, Hashmi et al. from the Pakistan National University of Science and Technology conducted a numerical simulation experiment on inter satellite adaptive optical communication in a laboratory. With moderate background noise and strong turbulence, the SR increased from 0.30 to 0.75 after adaptive optical correction [29].

In 2016, Pasupathi et al. from King Tanhaveer Engineering College in India conducted a wavefront detection adaptive correction experiment with a transmission distance of 940 m [30]. Using the SR of the imaging light spot as the evaluation index, the root mean square value of the wavefront phase decreased from 10^{1} before the correction to 10^{-6} after 1364 iterations.

In 2019, Carrizo et al. from the German Aerospace Center corrected the wavefront phase by iteratively updating the phase of a single focal plane speckle, thereby maximizing the power coupled into a single-mode fiber. Experimental results showed that the power increased by approximately 4 dB in less than 60 iterations of power measurements under strong turbulence conditions [31].

In 2020, Baykal et al. from Cankaya University in Turkey applied adaptive optics technology to a marine optical wireless communication system with pulse phase modulation to correct the distorted wavefront and reduce the bit error rate. The calculations showed that the ratio of the temperature to the salinity contributed less to the refractive index spectrum and mean square temperature dissipation rate, while the kinetic energy dissipation rate per unit mass of fluid increased, i.e., the ocean turbulence weakened, which led to more effective adaptive optical correction [32].

In 2020, Toselli et al. from the Fraunhofer Institute of Photonics in Germany theoretically studied the effectiveness of adaptive optical correction for Gaussian beams in ocean turbulence. When the first 15 modes were corrected and the transceiver aperture was sufficiently large to capture 4–5 speckles of the beam affected by ocean turbulence, the light intensity scintillation factor was reduced by seven times [33].

In 2020, Paillier et al. from the University of Paris Thackeray established an accurate numerical model for a space ground coherent optical link using adaptive optics technology and a digital phase-locked loop structure. It was a simple and reliable method for solving the carrier frequency and phase tracking problem in satellite ground coherent radio optical communication links [34].

In 2021, Ata et al. from the Austin University of Science and Technology in Turkey introduced adaptive optics technology to the field of underwater optical communication in natural turbulence. Theoretical calculations showed that even if adaptive optics technology was adopted, underwater optical wireless communication systems would still be seriously affected by regional and seasonal mean values and water refractive index fluctuations [35].

In 2021, Osborn et al. from Durham University in the United Kingdom studied a free space optical communication uplink using adaptive optics for pre compensation for Earth orbit and synchronous orbit satellite communication. The calculation results showed that, except for a small launch aperture and extremely low signal to noise ratio, adaptive optics could reduce the fading and bit error rate after pre-compensation for uplinks [36].

Since 1953, research on adaptive optics has expanded from the initial laser transmissions and astronomical observations to human eye imaging, inertial confinement fusion, laser processing, and many other aspects. The focus of research has shifted from astronomical imaging observations to laser communication. The evaluation index of the correction effect extends from a single optical index, such as the image resolution limit, wavefront peak valley value (PV), and wavefront root mean square value, to communication indexes, such as the coupling power and communication bit error rate.

Applying adaptive optics technology to the field of optical wireless communication has theoretically developed from initial simple numerical calculations [10, 11], such as tilt aberration correction [13], to today's complex numerical calculations, such as combinations with digital phase-locked loops [36], underwater optical communication wavefront correction [33], and satellite–ground laser communication uplink pre-correction [37, 38].

In terms of experiments, initial near-range wavefront correction [7, 30] and long distance near surface wavefront correction [21] were developed for experimental studies on satellite-based laser communication downlinks [19, 20] and uplinks [25]. For the transmission medium, the initial correction usually adopts the method of simulating turbulence [15, 16], then corrects measured atmospheric turbulence [26–28, 39], and gradually develops into the current marine turbulence wavefront correction [36].

Domestic Research Progress

In 1979, the "Infrared Adaptive-Optics Observation System for a 2.16-m Telescope," established by the Institute of Optoelectronics Technology of the Chinese Academy of Sciences and the Beijing Astronomical Observatory, realized horizontal and oblique atmospheric turbulence compensation [40–42].

In 2008, Yang theoretically analyzed an adaptive optical system using the stochastic parallel gradient descent (SPGD) algorithm. The gray value of the far-field spot increased from 20 before correction to 230 after correction. However, limited by the sampling frequency of the camera, only static or slowly changing aberrations could be corrected [43].

In 2008, Li et al. analyzed the problem of matching time and space bandwidths in adaptive optics systems. When $D/r_0 = 6$, the compensated light intensity reached approximately 50% of the diffraction limit, whereas when $D/r_0 = 2$, the diffraction limit reached approximately 90%. To obtain a satisfactory compensation effect in a strong turbulent atmosphere, the number of executing elements must be greatly increased [44].

In 2010, Xia et al. reduced the transmitting power by 50%, while maintaining the link bit-error rate at 1×10^{-6} under the same experimental conditions [45].

In 2011, Wu et al. corrected a wavefront aberration using an adaptive optics technique with an IM/DD communication mode. When using a multimode fiber (MMF) with a 62.5 μm core diameter, the bit-error rate decreased from 10^{-3} to 10^{-4} to below 10^{-6} after correction. When an SMF with a 10-μm core diameter was used, the deep fading of the receiving signal could be clearly observed, and the fading degree was related to the bit error rate. After an adaptive optics closed loop, the deep fading was suppressed and the bit-error rate of the system decreased from 10^{-3} to 10^{-12} [46].

In 2010, Han et al. used adaptive optics to simulate and analyze the distorted wavefront compensation for an atmospheric turbulence model with a gamma–gamma distribution. When correcting only the first five order Zernike coefficient, the bit-error rate was reduced from 10^{-1} before correction to 10^{-6} after correction [47].

In 2012, Wu et al. theoretically calculated that under the same turbulence condition, when pre calibrated adaptive optical correction was adopted at the transmitting end, compared with the receiving end, the received optical powers were 0.56 dBm and 0.34 dBm, and the fluctuation variances were 0.022 dBm and 0.042 dBm, respectively. The experimental results showed that the average bit error rate of the receiver decreased from 10^{-2} to 10^{-5} after adaptive optical correction. After adaptive optical correction at the receiving end, the average bit-error rate decreased from 10^{-3} before correction to 10^{-9} [48].

In 2013, Li et al. investigated the optimal adaptive optics closed loop bandwidth for wireless coherent optical communication links under strong turbulence conditions. When the number of photons detected per bit was greater than 100, the Greenwood frequency was sufficient to satisfy the servo bandwidth of the adaptive optical system. (The Greenwood frequency is used to represent the frequency

characteristics of turbulence. The RMS value of the phase power spectrum above that frequency on the distorted wavefront is 1 rad.) However, if the number of photons per bit was less than 70, the Greenwood frequency only satisfied the closed loop bandwidth under weak scintillation conditions. When the scintillation index was greater than 0.7, the servo bandwidth needed to be increased to at least twice the Greenwood frequency to satisfy the acceptable bit-error rate target [49, 50].

In 2013, Liu et al. calculated the aberration of laser signals. After correcting the aberration, the average efficiency of coupling spatial light to fiber improved from 10.32% to 63.82%, and the average bit-error rate of the optical wireless communication system decreased from 10^{-5} to 10^{-13}. As this result was only a simulation, the actual effect needs to be further verified [51].

In 2013, Wu et al. studied the influence of wavefront distortion on the bit-error rate of a binary phase-shift keying (BPSK) zero difference coherent detection system under different atmospheric turbulence conditions. The results showed that the aberration caused by atmospheric turbulence could directly affect the bit-error rate, independently of electrical noise. The PV value of the wavefront needed to be higher if the signal to noise ratio (SNR) decreased. When the PV value of the wavefront phase was 1.01λ, the mixing efficiency of the system was only 0.23. When the wavefront phase PV value was 0.53λ, the mixing efficiency of the system reached 0.65 [52].

In 2014, Liu et al. analyzed the influence of the atmospheric time-domain characteristics on the mixing efficiency and bit error rate of a wireless coherent optical communication system based on adaptive optics. The simulation results showed that the new wavefront sensing technology based on holography could adapt to weak atmospheric turbulence. The RMS value of the residual aberration was less than 0.04λ, the peak valley value was less than 0.25λ, and the fiber coupling efficiency was improved from nearly 30% to more than 70% [53].

In 2014, Li et al. studied an adaptive optical system with combined deformable mirrors to correct wavefront aberrations and improve the coupling efficiency. The simulation results showed that the coupling efficiency increased from the uncorrected 0.05% to 77.69% and 74.42% when corrected by two single deformable mirrors and 80.97% when jointly corrected [54].

In 2014, Liu et al. analyzed the influence of tilt and high-order aberrations on the mixing efficiency and bit error rate. The RMS value of the tilt aberration was used to estimate the change in mixing efficiency. First, the correctness of the theoretical analysis was verified via numerical simulation. Second, an experimental correction system for a two-stage fast-steering mirror and a 97-unit continuous curved-surface deformed mirror was designed. As the Greenwood frequency increased, a higher closed loop bandwidth was required to ensure the communication quality. When the Greenwood frequency was 160 Hz, the bit-error rate dropped below 10^{-9} in an adaptive optical system with a two stage fast mirror unit [55].

In 2014, Cao et al. used an SPGD algorithm to correct a wavefront distorted by atmospheric turbulence. They adopted a wavefront detection and control algorithm without a wavefront sensor to correct the wavefront, and the coupling efficiency of single-mode fiber was increased from 12% to more than 80% [56].

In 2014, Chao et al. analyzed the bit error rate performance of optical wireless coherent communication systems under different atmospheric turbulence conditions. The bit-error rate of the optical wireless coherent communication system was primarily affected by the directivity error under weak turbulence conditions. In the case of strong turbulence, the effect of high order aberrations on communication quality was increasingly serious, and the effect of correcting high order wavefront aberrations by the adaptive optics system was more obvious. With a 1550-nm wavelength as the communication band, the atmospheric turbulence intensity D/r_0 could be increased from 1 before correction to approximately 6.5 after adaptive optical correction. In addition, the mixing efficiency should be greater than 0.4, and the wavefront residual error after correction should be less than $\lambda/6$ [57–59].

In 2015, Huang et al. used an adaptive optical system to compensate for the wavefront distortion caused by atmospheric turbulence in a wireless coherent optical communication link. When the Fried's coherence length was $r_0 = 16$ cm, the amplitude fluctuation on the receiving aperture reduced the average SNR by about 24% [60, 61].

In 2015, Li et al. used adaptive optics to compensate for the influence of atmospheric turbulence. Under the same turbulence intensity conditions, a link channel at sea produced a larger bit-error rate than that on land [62, 63].

In 2015, Chen et al. used a 137-unit deformable mirror to compensate for the wavefront distortion caused by strong turbulence. When the phase RMS value was 0.3λ, the coupling efficiency decreased from 81% to 10% without phase distortion. In the case of strong turbulence, $D/r_0 = 15$, the SMF coupling efficiency of the system increased from 1.3% to 46.1% before and after correction [64].

In 2016, Zhao et al. combined adaptive optics technology with channel coding to correct phase distortions, and used wavefront correction and channel coding to further correct errors in each orbital angular momentum (OAM) mode. The simulation results showed that when the structural constant of the atmospheric refractive index was approximately $C_n^2 = 4.5 \times 10^{-14}\ m^{-2/3}$, the bit-error rate decreased from the original 10^{-1} to 10^{-3} after coding and wavefront distortion correction [65].

In 2018, Chen et al. reported that atmospheric turbulence had a serious impact on a wireless coherent optical communication system, based on an SMF coupling-system design of a 137 unit continuous curved-mirror adaptive optical unit. After adaptive optical compensation, the bit-error rate of the coherent communication was lower than 10^{-6}, which was significantly lower than the uncorrected bit error rate [66].

In 2018, Rui et al. developed multiple sets of adaptive optical ground stations of different calibers and designed the scale and arrangement form. They investigated the detection and correction module performance of an adaptive optical system according to the link characteristics and ensured the efficient coupling of a signal light to a single-mode fiber under strong atmospheric turbulence [67].

In 2019, Yang et al. investigated the performance improvement of wireless coherent optical communication systems using a large scale high speed adaptive optics system and a 349 unit continuous-surface deformable mirror. When the Greenwood frequency was 70 Hz, the bit-error rate of a 349 unit deformable mirror

adaptive optics closed loop was close to 10^{-9}, while the bit-error rate of the 97-unit deformable mirror only reached 10^{-6} [68].

In 2019, Chang et al. proposed a phase recovery algorithm correction method based on adaptive optics to compensate for the OAM beam distortion caused by ocean turbulence. Simulation results showed that in marine channels, adaptive optics technology based on a phase recovery algorithm could effectively compensate for the OAM beam distortion and improve the bit-error rate performance of the system [69].

In 2020, Rui et al. studied the statistical distribution of SMF coupling efficiency under adaptive optical compensation under different turbulence intensities. When the turbulence intensity $D/r_0 = 7$, the beam wavefront RMS value decreased from 0.42 to 0.04. Meanwhile, the coupling efficiency increased from 2.2% to 35.4%, on average. In addition, adaptive optical wavefront distortion correction improved the coupling efficiency of SMFs with a high probability [70].

In 2020, Chang et al. proposed an adaptive optics technique based on a phase recovery algorithm for single light intensity measurements to compensate for the wavefront distortion of OAM beams. The calculation results showed that the single-intensity measurement phase recovery algorithm was more robust than a double-intensity measurement phase recovery algorithm [71].

In 2020, Gu et al. combined a convolutional neural network with an SPGD algorithm to improve the coupling efficiency of spatial light to fibers. The calculation results showed that when the Strehl ratio was 0.9, the SPGD algorithm required several hundred iterations, whereas the algorithm combining a convolutional neural network and SPGD algorithm required a dozen or less iterations. The convergence time required by the SPGD algorithm was almost eight times that of the hybrid algorithm [72].

In 2021, Jiang applied adaptive optics technology to improve the coupling efficiency of a single-mode fiber on a 1 km experimental link. The experimental results showed that the coupling efficiency of the SMF increased from 5% in the open loop state of adaptive optics to 10–25% in the closed loop state [73].

In 2021, to avoid the wavefront distortion caused by the turbulence amplitude exceeding the measurement range of a Shack–Hartmann wavefront sensor, Zhang et al. proposed a method combining a standard wavefront sensor and a Shack–Hartmann wavefront sensor to expand the detection range. After correction, the communication bit-error rate decreased from 10^{-1} to 3×10^{-6}, and the correlation of the dot plot increased from 0.6 to 0.94 [74].

In China, research on adaptive optics technology is usually based on large projects; however, comprehensive and in-depth research on independent adaptive optics technology is lacking. Moreover, the application of adaptive optics technology to optical wireless communication is typically based on numerical calculation simulations. Numerical calculations mainly involve fiber coupling [56], closed loop bandwidth [49, 50], pre-correction [48], multi-corrector combinations [54], and ocean turbulence [69]. Unfortunately, these are limited by many conditions, such as experimental equipment and experimental sites.

Only a few scientific-research institutions have carried out relevant experiments, such as the Institute of Optoelectronic Technology, Chinese Academy of Sciences [58, 59, 64, 75]. Moreover, the experimental scale lags behind that of developed countries [76]. In addition, the uncertainty of atmospheric turbulence, the complexity of the system itself, the precision of signal capture and tracking, and the duration of research are all important factors restricting the development of adaptive optics in China.

Adaptive Optics in the Optical Wireless Communication Field

In recent years, Ke Xizheng et al. of Xi'an University of Technology have studied the following:

1. Laser transmission characteristics in atmospheric turbulence [77, 78]
2. Encoding and decoding of wireless laser communication systems [79]
3. Optical wireless subcarrier modulation
4. Channel estimation and equalization [80]
5. Alignment tracking and acquisition technology
6. Heterodyne coherent detection technology
7. Optical transceiver system design
8. Optical orthogonal frequency-division multiplexing modulation
9. Optical multi-input multi-output systems [81, 82]
10. Optical wireless transmission theory and technology under complex environments and systems [83]

Other key technologies of space optical communication have been deeply studied theoretically and experimentally verified, and gratifying results have been achieved [84, 85].

In 2016, a free-space optical (FSO) communication system terminal based on the IM/DD modulation and detection mode was successfully developed. The entire machine weighed less than 8 kg. The terminal had standard Ethernet data access capability and could reliably transmit duplex data, voice, and images.

Since 2017, Ke et al. have studied an optical wireless communication system for ultra-long distance transmission [86] and conducted experiments on a differential binary phase-shift keying (DBPSK) modulation coherent detection system [87–92].

Figure 23.1 shows the 100-km optical wireless communication experimental link of Xi'an University of Technology, and Fig. 23.2 shows a photo of the coherent optical communication receiving antenna. The optical wireless coherence communication experimental link based on coherent detection is located at both ends of Qinghai Lake. The transmitting end is located in Jiannan Mountain, Erlang, Qinghai Lake, and the receiving end is located in Quanji Township, Qinghai Lake. The transmitting terminal is composed of a laser source, electro-optic modulator, fiber amplifier, transmitting antenna, coarse-aiming platform, and micro-aiming platform. The transmitting antenna is a Kepler optical transmission antenna.

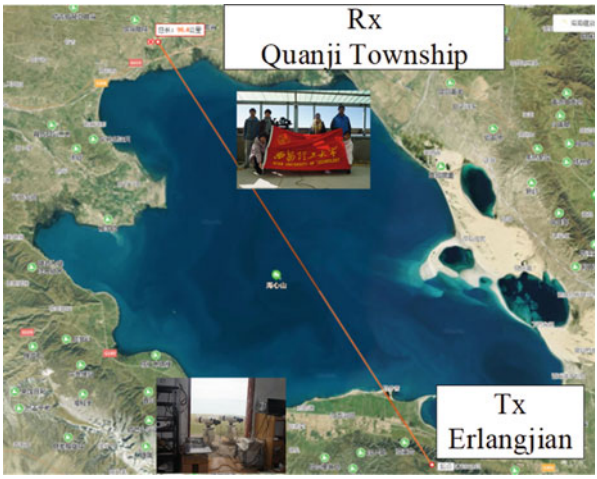


Fig. 23.1 Diagram of the 100-km optical wireless communication experimental link of Xi'an University of Technology

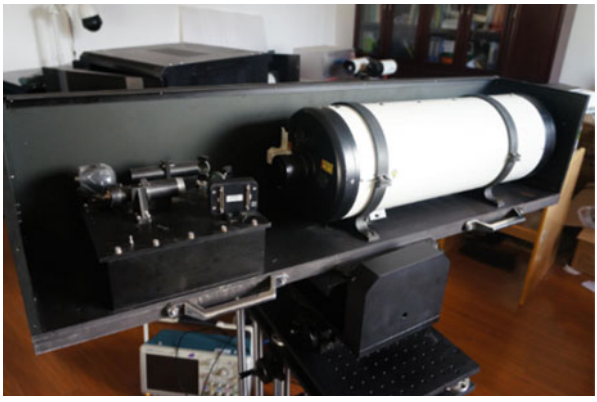


Fig. 23.2 Internal structure of the wireless coherent optical communication receiving antenna

The coarse-aiming platform, micro-aiming platform, and optical-antenna focusing system constitute a composite axis continuous focal length adjustment structure to enable beam alignment and tracking. The receiving end is mainly composed of a Cassegrain reflector receiving antenna, adaptive optical system, and coherent detection system.

An adaptive optical system composed of a bias mirror, deformable mirror, and wavefront sensor is used to correct the optical signals. After the wavefront correction, the laser is coupled to the mixer and balance detector by a fiber, and the signal is recovered after mixing with the local oscillator light.

After a long distance transmission of the light front, the light spot is completely broken, owing to atmospheric turbulence over Qinghai Lake. The light flicker caused by strong turbulence makes the sampling point incomplete and leads to difficulty in wavefront reconstruction. This directly leads to non continual data

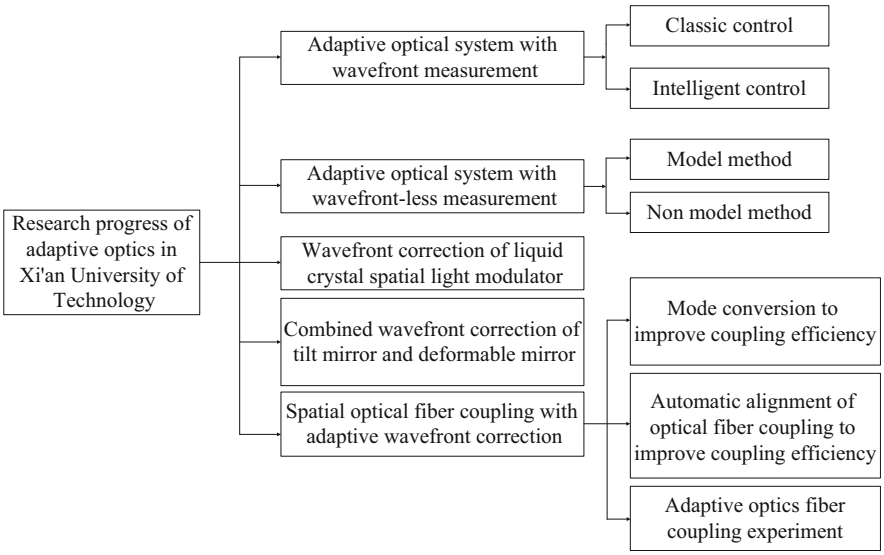


Fig. 23.3 Content structure of adaptive optics research

acquisition. The phase distortion of the wavefront is large, and the rate of change is high. Based on the collected wavefront data, the atmospheric coherence length was calculated to be 1.3 cm using the Hartmann wavefront measurement method.

When the transmitter uses 0 dBm power for beam emission, the power of the receiver coupling into the single-mode fiber is about −40.33 dBm, and its isolation is approximately 40 dB. With an increase in the output power of the fiber amplifier, the power coupled into the fiber increases, and the coupled light power of the double emission is higher than that of the single emission. Moreover, under the same power output, the variance in the double-emission coupled power is smaller than that of the single-emission coupled power. This shows that multiple structures can effectively suppress the light intensity flicker and coupling power fluctuation caused by atmospheric turbulence.

Ke et al. from Xi'an University of Technology have conducted studies on an adaptive optical system with wavefront measurement, adaptive optical system without wavefront measurement, wavefront correction of a liquid crystal spatial light modulator, wavefront correction of a combined tilt mirror and deformable mirror, and adaptive wavefront correction of spatial light fiber coupling. The structure of the study is shown in Fig. 23.3.

Adaptive Optical System with Wavefront Measurement

Adaptive optical systems can be divided into two types: with and without wavefront measurement. Figure 23.4 shows the principle diagram of an adaptive optical system

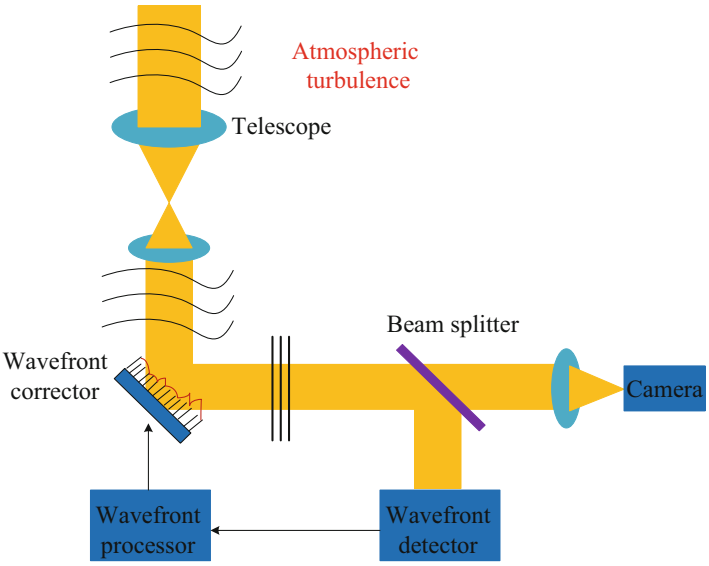


Fig. 23.4 Schematic diagram of an adaptive optical system

with wavefront measurement. The wavefront sensor measures the wavefront error in real time and transmits the measured wavefront error to the wavefront controller. The wavefront controller receives the distorted beam information from the wavefront sensor and calculates and obtains the control voltage of the wavefront corrector using an algorithm. The wavefront corrector receives the control voltage from the wavefront controller, corrects the wavefront distortion, and improves the communication quality.

Figure 23.5 shows the optical path of the wavefront adaptive optical system designed by Xi'an University of Technology.

As a traditional calibration method, the push pull method applies push and pull commands to the deformable mirror to complete the response matrix calculation [93]. Figure 23.6 shows the surface pattern generated by the deformable mirror driven by the 10th driver and the slope pattern collected by the wavefront sensor during the push pull method to solve the response matrix. Figure 23.7a shows the response matrix calculated using the traditional push pull method.

A Hadamard matrix was used to construct the transmission-voltage matrix and act on the deformation mirror. Several drivers were virtualized to realize the response-matrix calculation based on the Hadamard matrix [94]. Figure 23.7b shows the response matrix calculated using the Hadamard transformation method.

A prediction model of the adaptive optical system was established according to the linear subspace system identification method [95]. The DC gain of the adaptive optical system was obtained according to the control theory, namely, the response matrix of the system. Figure 23.7c shows the response matrix calculated using the system identification method.

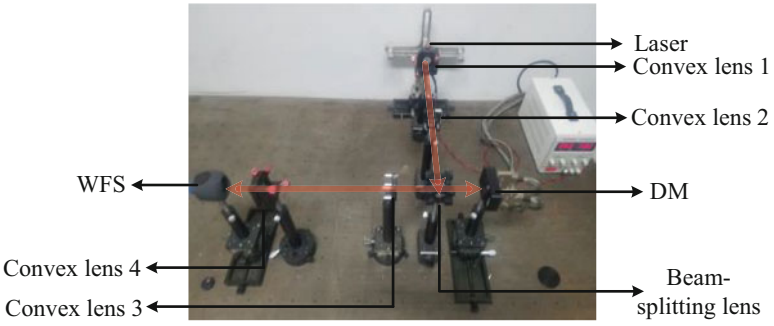


Fig. 23.5 Structure diagram of a wavefront adaptive optical system

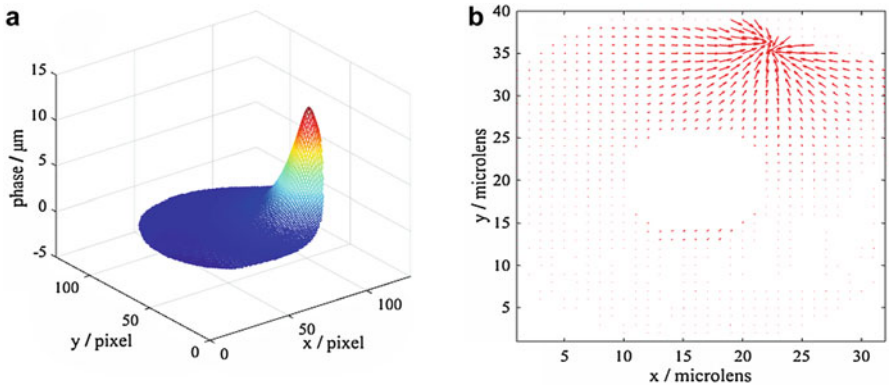


Fig. 23.6 (**a**) Deformable mirror profile; (**b**) wavefront-slope diagram driven by the 10th driver of the push–pull method

Classic Control

The proportional–integral–differential (PID) control algorithm is the most widely used algorithm for adaptive optical closed loop control. Figure 23.8 shows the principle diagram of the PID control algorithm for an adaptive optical system, where S_{tar} represents the target slope, S_{mea} the acquisition slope, S_{err} the error slope, V_{out} the output voltage of the deformable mirror, and $CM_{errslope2cmd}$ represents the command matrix from slope to voltage.

After the error slope is transformed into the error voltage by the command matrix, the voltage calculated by the PID algorithm is applied to the deformation mirror to correct the distorted wavefront. The wavefront sensor collects the wavefront slope again to complete the iterative closed loop control. Figures 23.9 and 23.10 show the wavefront PV curve and phase diagram of wavefront phase correction when the k_p and k_i parameters are changed, respectively. Compared with the initial state, the corrected wavefront is very close to the plane wave shape, and the concavity and surface roughness are greatly reduced.

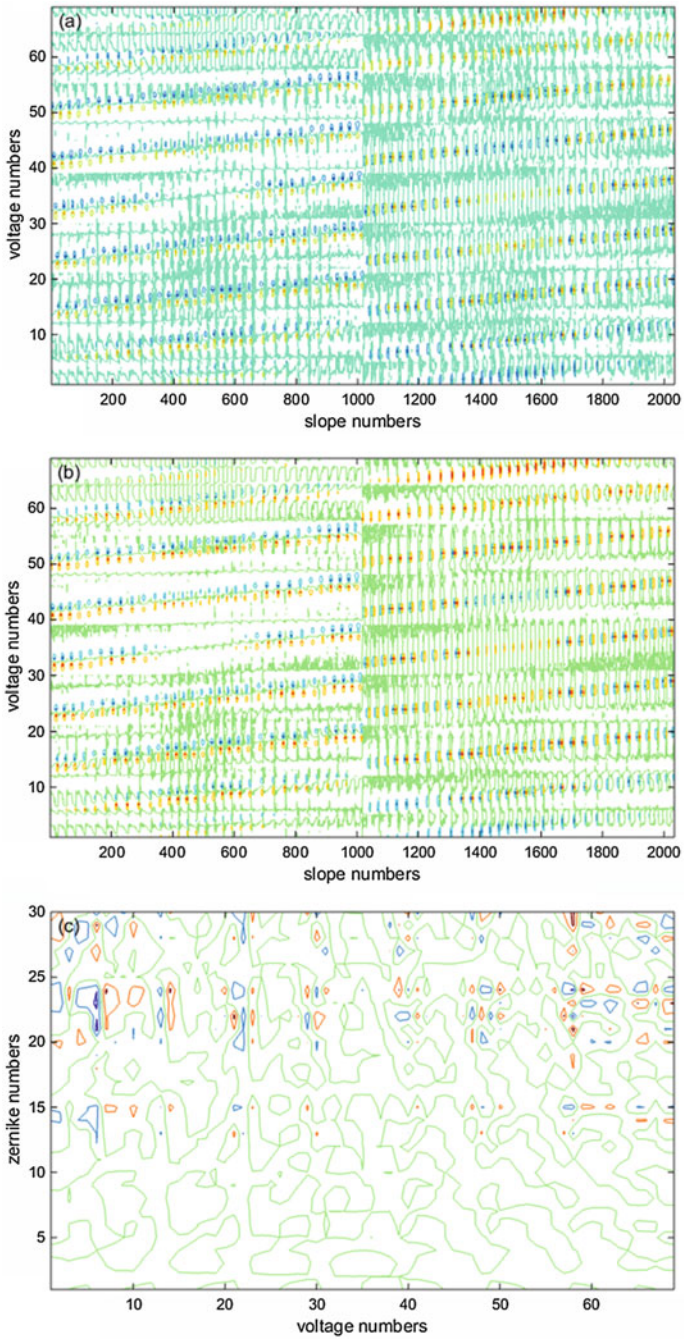


Fig. 23.7 Response matrices. (**a**) Traditional push–pull method; (**b**) Hadamard method; (**c**) system identification method

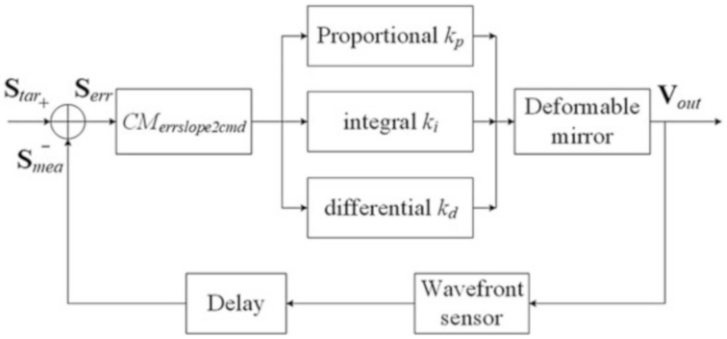


Fig. 23.8 Principle diagram of an adaptive optical PID control algorithm

When the response matrix (RM) from the deformable mirror voltage to the wavefront sensor slope is not limited, the RM between any quantity generated by the deformable mirror and any quantity collected by the wavefront sensor is established, and PID control is used to realize the closed loop control of the wavefront adaptive optical system.

Intelligent Control

Because the adaptive optical system is a typical multi-input multi-output (MIMO) linear system, its closed loop control can be realized by solving approximate solutions of linear equations using iterative arithmetic. Jacobi and Gauss–Seidel iterations are commonly used iterative methods for solving linear equations. The correction effect of the Gauss–Seidel algorithm in the field of adaptive optics is controlled by the accuracy of the iteration error. Figure 23.11 shows the wavefront phase diagram after correction with different convergence values. Selecting a suitable convergence value is the key step of the Gauss–Seidel iterative algorithm for realizing closed loop control.

Iterative learning control (ILC) constructs an iterative relationship from the linear relationship between the wavefront Zernike polynomial and the control voltage of the corrector. When an appropriate γ value is selected, it can make the iterative control algorithm converge to the optimal effect [97, 98]. Figures 23.12 and 23.13 respectively show the wavefront closed loop convergence curve and wavefront phase diagram using the ILC algorithm under different parameters. The greatest advantage of this algorithm is that it has a fast convergence rate when the parameters are appropriately selected.

Using the PID algorithm, Ke et al. from Xi'an University of Technology designed a double fuzzy adaptive PID control for an adaptive optical system [99]. Fuzzy control consists of four parts: a fuzzer, rule base, fuzzy inference machine, and fuzzer solution. Figure 23.14 shows the PV convergence curve of the wavefront with fuzzy control wavefront correction under different parameters.

Simulations and experiments showed that the fuzzy PID method effectively corrected the distorted wavefront signal, improved the system response speed,

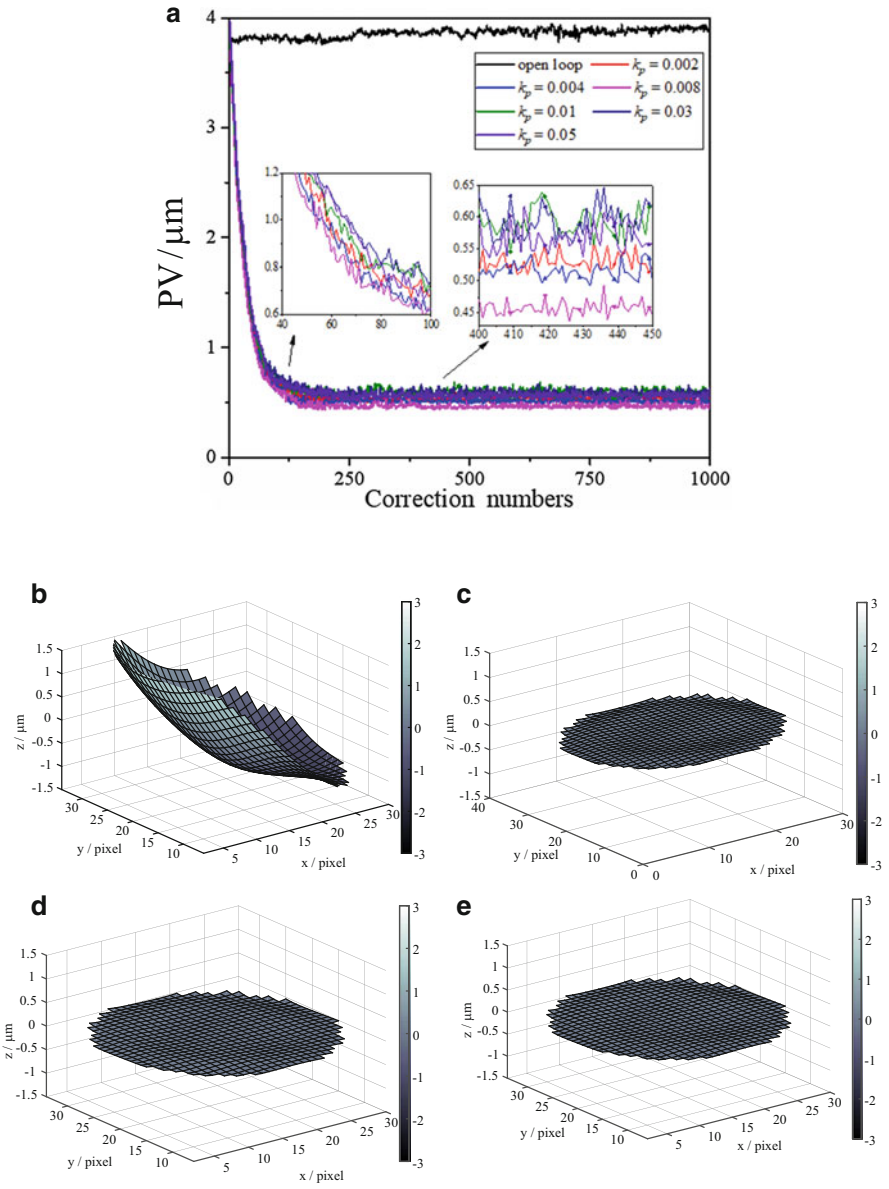


Fig. 23.9 Wavefront correction. (**a**) Wavefront PV curve; (**b**) open loop state under different k_p parameters; (**c**) $k_p = 0.002$; (**d**) $k_p = 0.008$; (**e**) $k_p = 0.03$ [96]

reduced the parameter-adjustment difficulty, and met the wavefront correction requirements for laser communication.

After the laser is transmitted over a long distance, the wavefront distortion caused by strong turbulence renders the wavefront sensor unable to accurately capture the

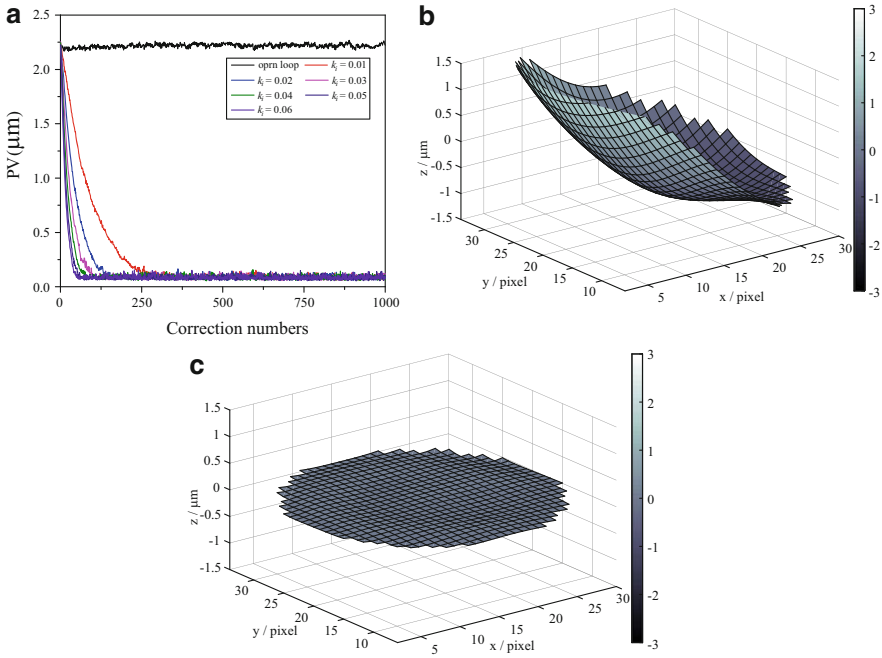


Fig. 23.10 Wavefront correction with different k_i parameters. (**a**) Wavefront PV curve; (**b**) open loop state; (**c**) $k_i = 0.01$ [96]

wavefront information in real time. The deformable mirror execution lags, which leads to errors in wavefront correction and decreases its accuracy. Based on the PID algorithm, the data are predicted, and a wavefront prediction model is established to reduce the servo delay and improve the correction rate.

Figure 23.15 shows the convergence curve of the adaptive optical wavefront prediction and correction. The closed loop convergence speed of the prediction algorithm is fast, and the correction speed of the system is relatively fast. Before the closed loop system stabilizes, the correction effect of the prediction algorithm is more evident than that of the PID algorithm. When the system reaches a stable state, the correction effects of the two algorithms are essentially the same.

Adaptive Optical System Without Wavefront Measurement

Based on wavefront measurements, Ke et al. from Xi'an University of Technology conducted research on an adaptive optical system with wavefront-free measurement [101]. Wavefront-free sensing technology does not require a wavefront sensor to measure and reconstruct the wavefront [102, 103] to obtain the wavefront distortion information. Rather, it directly converts the distortion information into performance indicators, such as the light intensity, SR, and coupling efficiency.

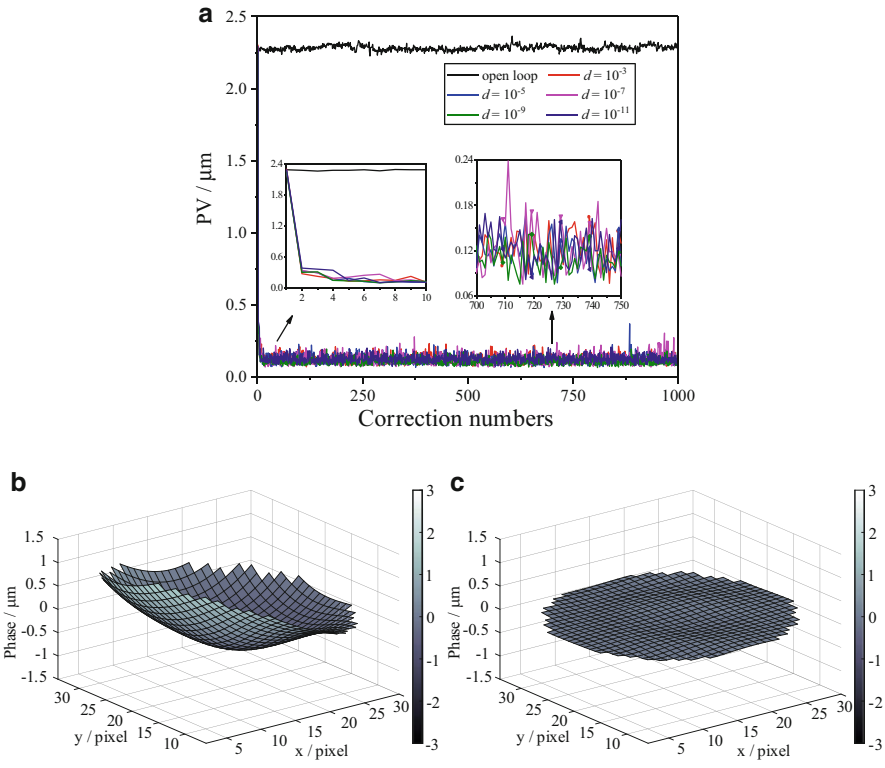


Fig. 23.11 Wavefront phase after correction with different convergence values δ. (**a**) PV value; (**b**) initial moment; (**c**) δ = 10–11 [96]

An optimization algorithm was used to iteratively optimize the performance indicators and control the constant deformation of the corrector. Thus, the distorted wavefront was corrected [104]. The wavefront-free adaptive optical system is shown in Fig. 23.16. LSM refers to liquid crystal spatial light modulator, CCD refers to charge coupled device, and DM refers to deformable mirror.

Mode Methods

The mode method holds that a distorted wavefront can be represented by a specific mode [105]. The algorithms are generally divided into two types.

One selects a mode to simulate the distorted wavefront and determines the transformation matrix between the mode coefficient and the influence function of each actuator of the deformation mirror. Then, combined with an intelligent optimization algorithm, the required mode correction coefficient is iteratively calculated to correct the distorted wavefront.

The other type calculates the performance index as a system function according to the detected light spot and uses the relation between the function and the mode

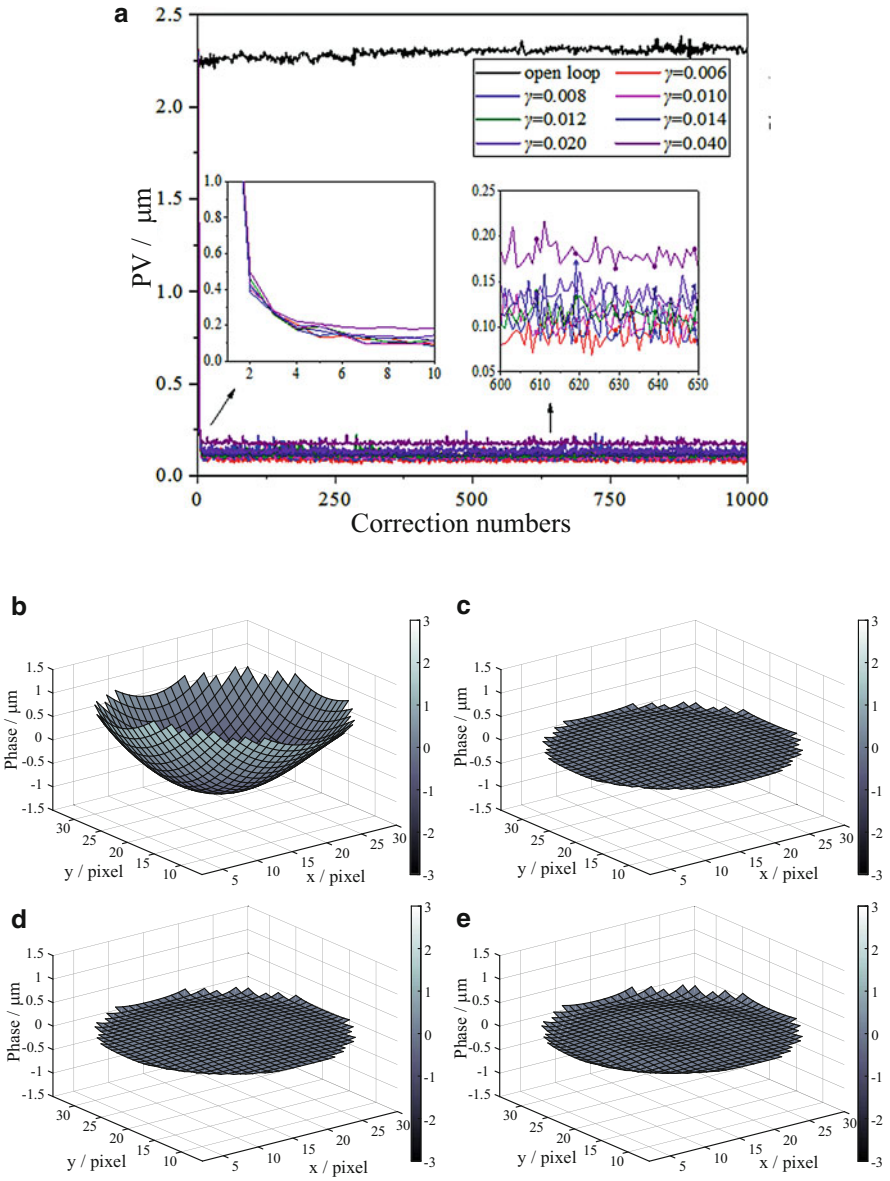


Fig. 23.12 Different γ-value corrections for the wavefront phase. (**a**) PV value; (**b**) open loop time; (**c**) $\gamma = 0.006$; (**d**) $\gamma = 0.014$; (**e**) $\gamma = 0.04$ [96]

coefficient to calculate the mode correction coefficient needed to correct the wavefront error. Because the selected base mode needs to satisfy a reciprocal orthogonal relationship, commonly used base modes include Lukosz modes and eigenmodes [106].

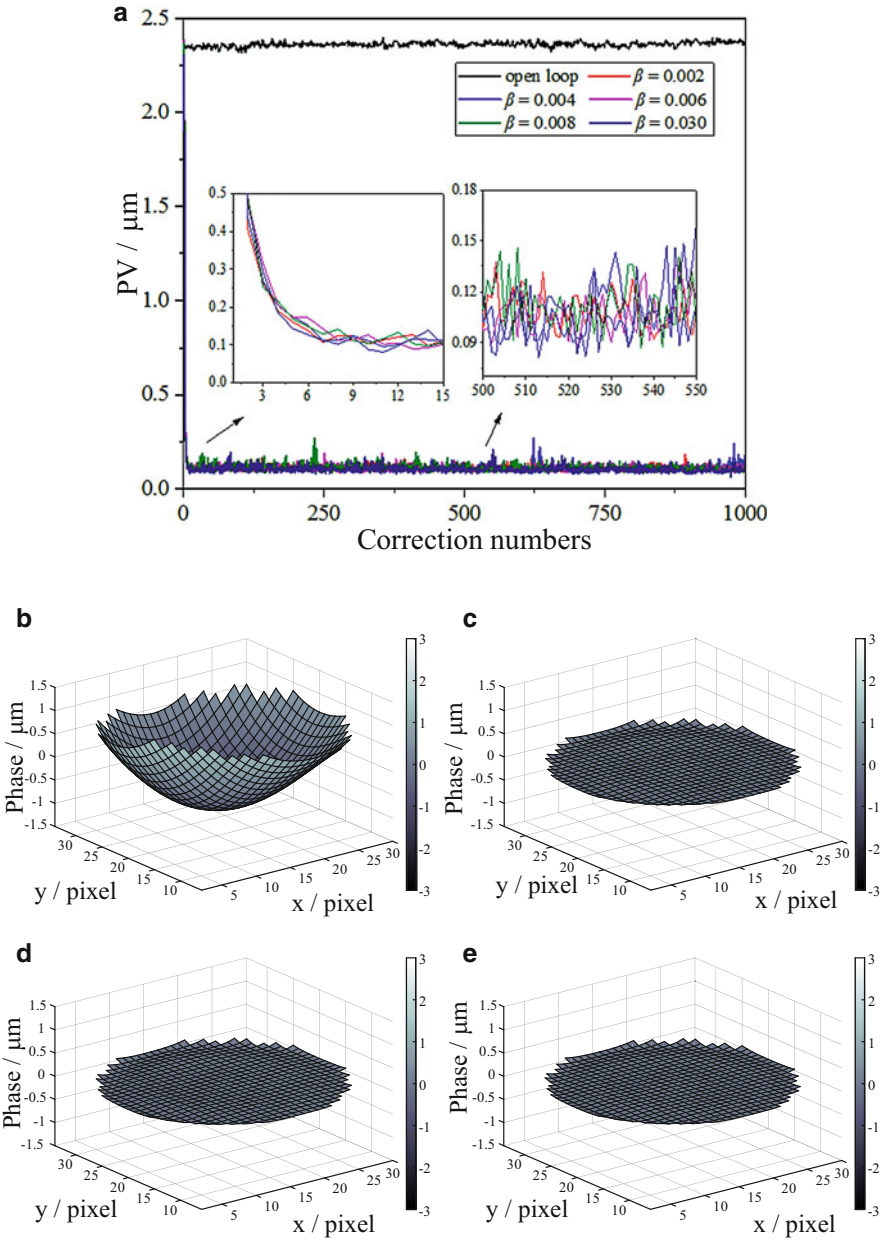


Fig. 23.13 Different β-value corrections for the system wavefront phase. (**a**) PV value; (**b**) open loop time; (**c**) β = 0.002; (**d**) β = 0.008; (**e**) β = 0.03 [96]

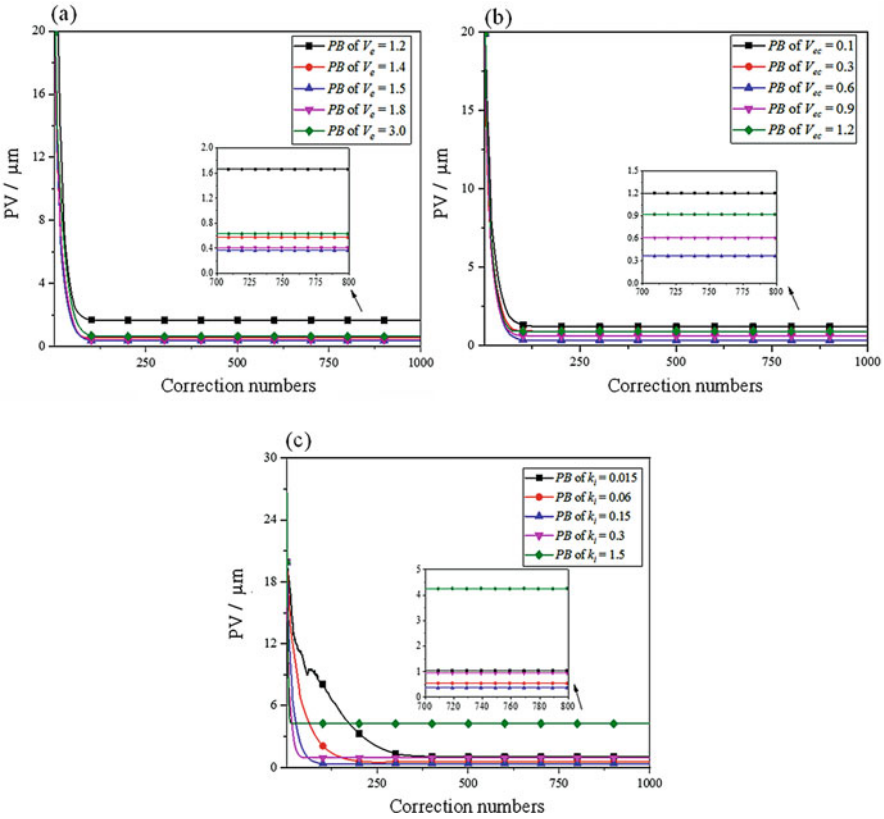


Fig. 23.14 Wavefront PV curve for fuzzy control correction [100]. (**a**) Different V_e; (**b**) different V_{ec}; (**c**) different k_i

The wavefront free correction algorithm for the eigenmode of a deformable mirror first establishes an optical system model to simulate the transmission of collimated Gaussian beams in the atmosphere and generate distorted light spots. Subsequently, it establishes a deformation mirror model and deduces the corresponding deformation mirror eigenmode as the base model. Finally, the system function is calculated and the wavefront aberration is corrected according to the principle of the eigenmode method of the deformable mirror.

Figure 23.17 shows the wavefront correction curve using the deformation mirror eigenmode method, and Fig. 23.18 shows the pattern of light spots before and after correction. The light intensity gradually increases with an increase in the number of iterations.

Non-mode Methods

The working principle of the adaptive optical non-mode method is to obtain the driving voltage of the wavefront corrector according to the optimization algorithm,

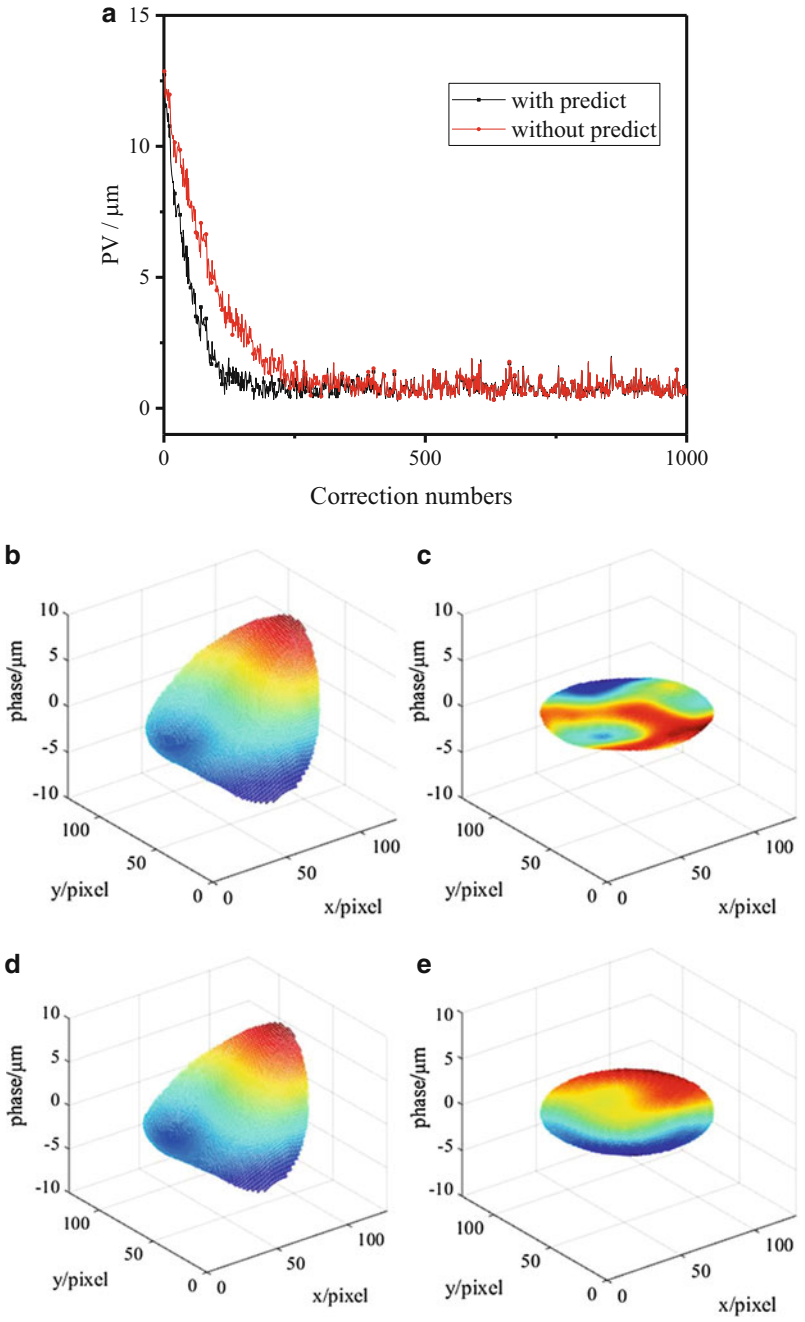


Fig. 23.15 Adaptive optical wavefront prediction correction. (**a**) Convergence curve; (**b**) unpredicted open loop phase; (**c**) unpredicted closed loop phase; (**d**) predicted open loop phase; (**e**) predicted closed loop phase



Fig. 23.16 Photo of the wavefront-free adaptive optical system

so that the objective function can reach the maximum value. Commonly used algorithms include gradient-descent algorithms, such as the conjugate gradient method and SPGD algorithm [108, 109], and heuristic algorithms, such as the simulated annealing (SA) [110], mountain climbing, and genetic algorithms [111].

Based on the light intensity distribution, the SPGD algorithm continuously optimizes to find the optimal value of the control voltage of the wavefront corrector so that the value of the objective function can meet the requirements [112]. From the numerical simulation perspective, wavefront distortion correction of a wavefront detection adaptive optics system based on the SPGD algorithm is studied, and the influence of different parameter values in the SPGD algorithm on the correction effect is analyzed [113].

Figure 23.19 shows the correction curve and light spots before and after correction using the SPGD algorithm. Before correction, far-field light spots have serious speculation and energy nonconvergence.

The SPGD algorithm takes 1550 nm optical power as feedback and uses fiber coupling, a deformable mirror, and a processor to form a loop to correct the 1550 nm optical power. The corrected optical power changes in the fiber coupling output are shown in Fig. 23.20.

After correction, the optical power output of the signal from the wavefront fiber tail increases obviously, and the coupling efficiency increases by approximately 32 times. The improvement in the coupling efficiency directly contributes to the communication performance. However, this method is only applicable to a weak-turbulent static wavefront environment and is not applicable to the coupling efficiency fluctuation caused by a strong turbulent environment.

Ke et al. from Xi'an University of Technology carried out wavefront correction experiments on vortex optical adaptive optics with non-mode and no wavefront methods. The Gerchberg–Saxton algorithm for phase recovery is an indirect measurement method used to reconstruct the wavefront phase from the wavefront intensity and inversely deduce the phase information from the optical intensity distribution.

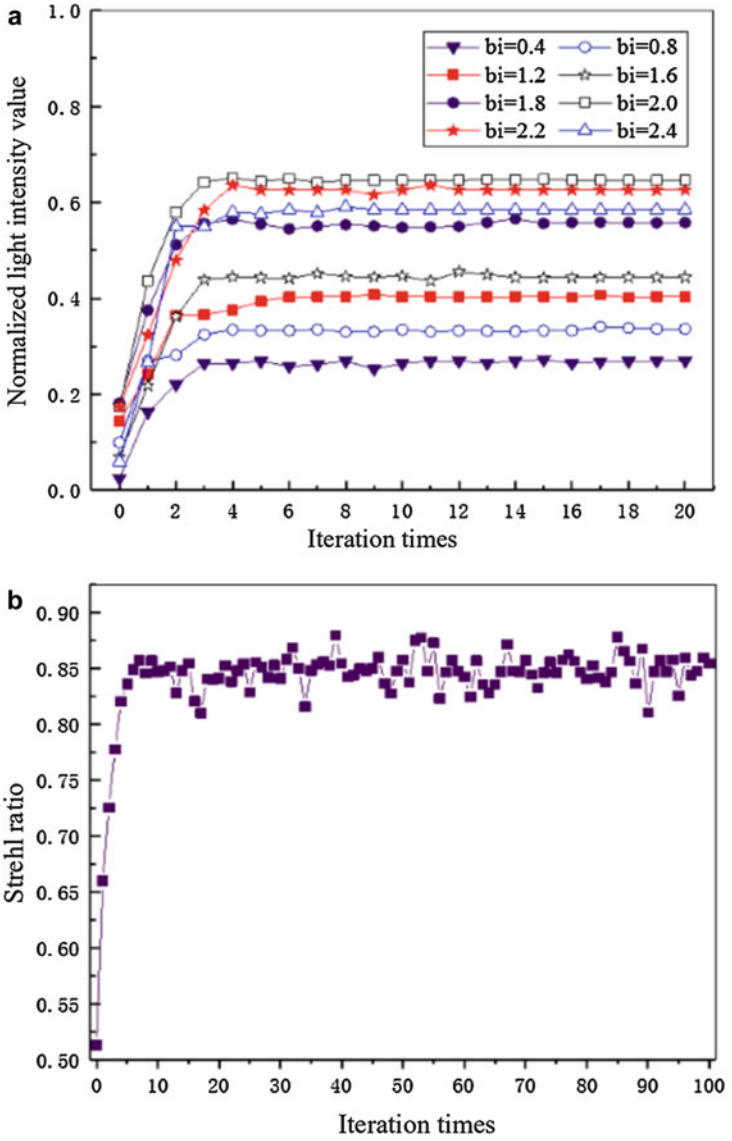


Fig. 23.17 Variation curve of the SR intensity corrected by the eigenmode of the deformable mirror. (**a**) Simulation; (**b**) experiment [107]

The idea of the algorithm is as follows: Based on the light intensity distribution at the incident plane and the output plane, the phase distribution information of the required light field can be obtained by controlling the initial conditions and performing iterative operations [116].

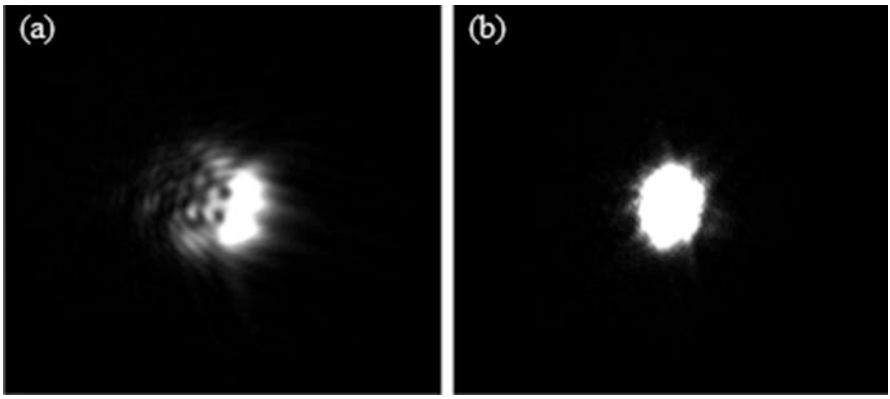


Fig. 23.18 Light spots before and after wavefront correction. (**a**) Before correction; (**b**) after correction [107]

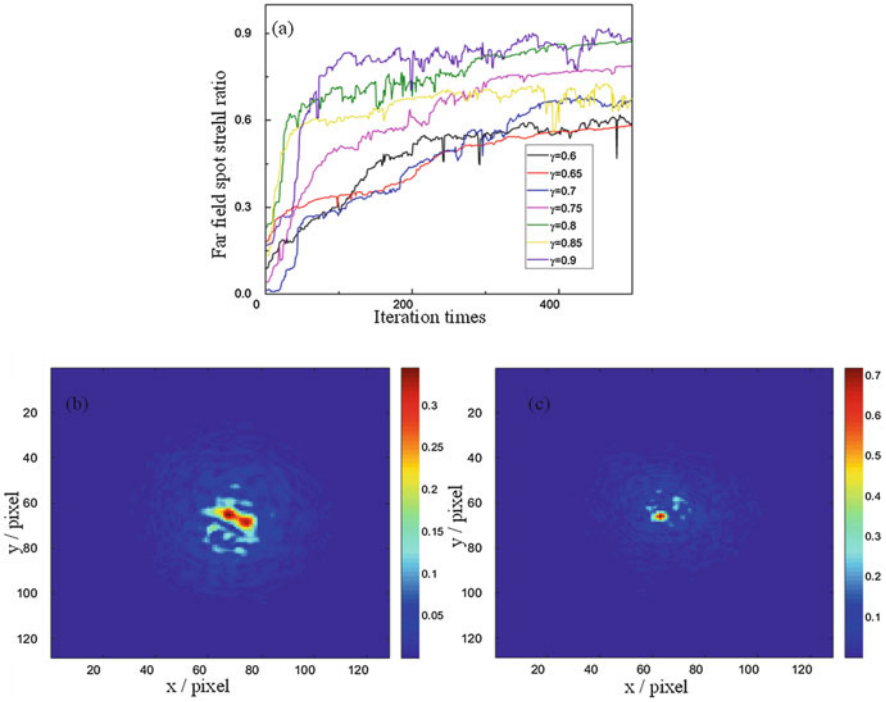


Fig. 23.19 Wavefront correction based on the SPGD algorithm [114]. (**a**) Variation curve of the light intensity SR; (**b**) spot before correction; (**c**) spot after correction

Figure 23.21 shows the intensity distribution of each order of vortex beam before and after correction using the Gerchberg–Saxton phase recovery algorithm. After the correction, the toroidal distribution of the single-mode vortex beam after turbulence

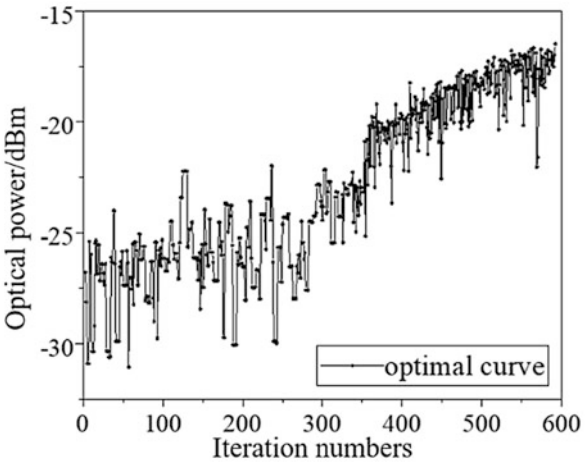


Fig. 23.20 Fiber coupling correction curve of the wavefront-free correction method [115]

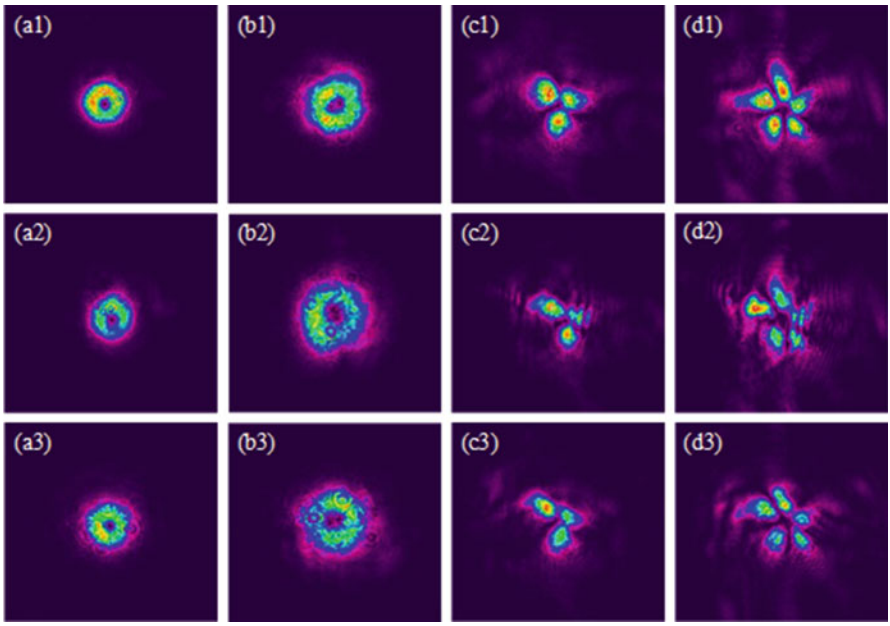


Fig. 23.21 Intensity distribution of vortex beams of each order before and after correction using the Gerchberg–Saxton phase recovery method [117]. (**a1**)–(**d1**) no turbulence; (**a2**)–(**d2**) turbulent; (**a3**)–(**d3**) after correction

is restored to a large extent, and the energy distribution on the toroidal spot becomes uniform. As can be seen from Fig. 23.21, the Gerchberg–Saxton phase recovery algorithm can effectively correct wavefront distortion for both single-mode and multimode multiplexed vortex beams.

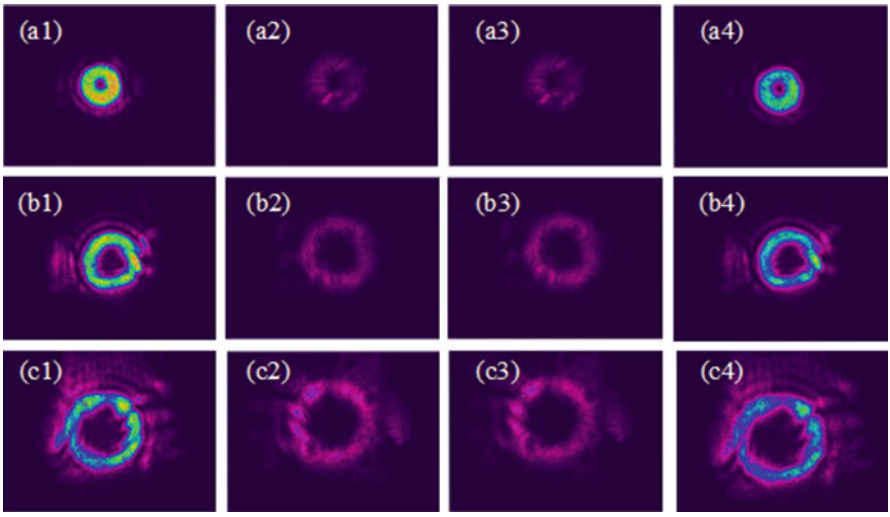


Fig. 23.22 Intensity distribution of a single vortex beam corrected by the phase difference method [119]. (**a1**)–(**c1**) initial light intensity; (**a2**)–(**c2**) light intensity at the focal plane; (**a3**)–(**c3**) light intensity at the defocusing plane; (**a4**)–(**c4**) light intensity after correction

The phase difference (PD) method is a wavefront-free detection technique that indirectly inverts the phase distribution of the wavefront from the light intensity information of the image surface of the imaging system. In this algorithm, the intensity information of the focal and defocusing planes of the system is collected, and the objective function is established according to the point diffusion function of the two channels. Finally, the phase distribution of the wavefront on the pupil plane of the imaging system is obtained by solving the optimization function.

Figure 23.22 shows the intensity distribution of a single vortex beam corrected and compensated by a wavefront correction system using the phase difference method [118]. After the atmospheric turbulence disturbance, the beam is distorted, and the intensity distributions of the focal plane and defocusing plane become dispersed, compared with the initial state. After correction using the phase difference method, the distribution of the light intensity is more uniform than the distortion moment, and the energy on the ring spot is enhanced. The phase difference method can effectively correct the wavefront distortion of a single vortex beam.

Wavefront Correction Using Liquid Crystal Spatial Light Modulators

Liquid crystal spatial light modulators (LC-SLM) are based on the electronically controlled birefringence principle of a liquid crystal. They can continuously modulate the light wave phase by changing the electric driving signal of the liquid crystal. Figure 23.23 shows a block diagram of a liquid crystal spatial light-front correction system based on a coherent optical communication system.

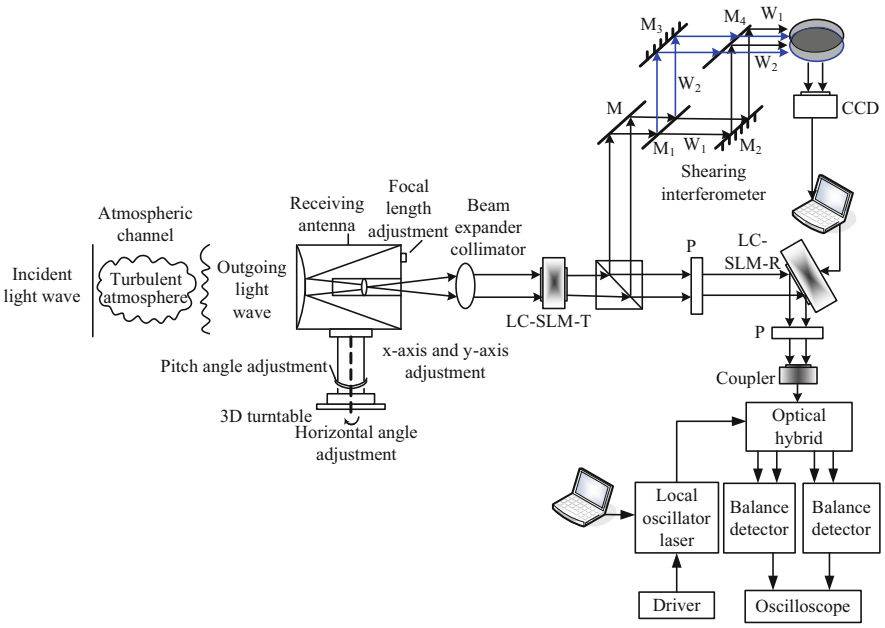


Fig. 23.23 Schematic diagram of LC-SLM wavefront correction for a wireless coherent optical communication system [120]

Figure 23.24 shows the voltage value of the intermediate frequency (IF) signal output by the oscilloscope without wavefront correction. Figure 23.24a is the voltage amplitude of the IF signal in the same phase direction (I channel), and Fig. 23.24b is the voltage amplitude of the IF signal in the orthogonal direction (Q channel). After correction, the voltage amplitude of the IF signal output by the balanced detector tends to stabilize. Although the voltage amplitude of the IF signal fluctuates slightly with the change in time in the experiment, there is no great difference between the output sine wave amplitudes, which is basically stable.

Compared with the voltage amplitude of the IF signal before correction in Fig. 23.24, the voltage amplitude of the IF signal is increased by a factor of approximately 1.7 times after wavefront correction by LC-SLM, and the SNR after correction is increased by approximately three times [121].

Combined Wavefront Correction Using Tilt and Deformable Mirrors

Introducing a tilt mirror (TM) [123] or a multistage deformable mirror (DM) for wavefront correction is the current development trend of adaptive optical systems for large-distortion wavefront correction. A tilt mirror or a large-stroke deformed mirror (usually called a woofer) is used to correct low-order wavefront components, and a small-stroke deformed mirror (usually called a tweeter) is used to correct high-order

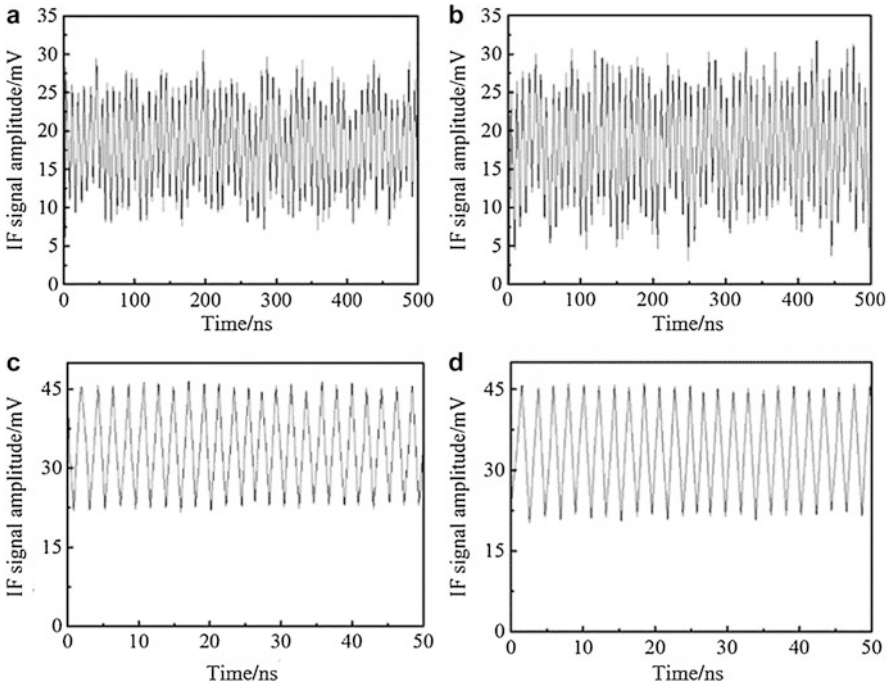


Fig. 23.24 Amplitude of the IF signal before and after correction [122]. (**a**) I channel before correction; (**b**) Q path before correction; (**c**) I path after correction; (**d**) Q circuit after correction

wavefront components. Based on the orthogonality of the Zernike polynomials, the space of the correction region between them is clearly divided [124].

A common multi corrector adaptive optical system includes a piezoelectric biasing TM, based on two-dimensional motion, and a DM69 deformable mirror with an independent element. This invention relates to an adaptive optical system consisting of a large-stroke DM69 deformable mirror with few independent elements and a small-stroke DM292 deformable mirror with many independent elements. Because the laser communication system has a low requirement for wavefront correction accuracy, and the cost of a DM292 is higher than that of a TM, in practice, in an adaptive optical system, a TM and DM69 are typically used for closed loop correction for system communication.

Figures 23.25 and 23.26 show the wavefront PV curve and phase diagram for TM-alone correction, DM69-alone correction, and simultaneous TM and DM69 correction, respectively. As shown in Fig. 23.25, the simultaneous correction effect of the TM and DM69 closed loops is better than that of TM or DM69 alone. The introduction of the TM alleviates the large-travel problem of the DM actuator caused by excessive wavefront distortion, and the power of the deformable mirror can be alleviated when correcting a wavefront phase with the same distortion.

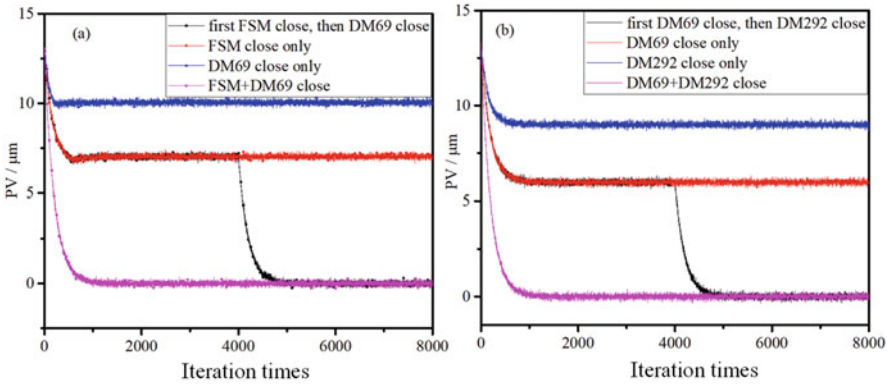


Fig. 23.25 PV convergence curve of multi-corrector wavefront correction. (**a**) TM + DM69; (**b**) DM69 + DM292

Spatial Optical Fiber-Coupled Adaptive Optical Wavefront Correction

Automatic Fiber Coupling Alignment

To improve the efficiency of coupling space light to fiber, Ke et al., Xi'an University of Technology, designed a fiber coupling system based on the automatic alignment of two-dimensional piezoelectric ceramics [125]. By changing the voltage applied to the piezoelectric ceramics, the fiber is displaced and aligns the optical axis [126].

Figure 23.27a shows the variation curve of the X-axis output voltage after a simulated annealing algorithm is adopted. It can be seen from the figure that the curve always fluctuates at the beginning, i.e., the entire system is looking for the maximum value. Then, it changes to a stable state, i.e., the optimal coupling position is found. At this time, the voltage value shifts up and down at 23.68 V and the amplitude does not exceed 0.1 V. At this time, the mean voltage is 23.65 V and the variance is 0.04432 V.

Figure 23.27b shows the variation curve of the Y-axis output voltage after the simulated annealing algorithm is adopted. It can be seen from the figure that the curve always fluctuates at the beginning, i.e., the entire system is looking for the maximum value. The fluctuation indicates the process of searching for the best coupling position, and then, it is in a stable state, i.e., the best coupling position is found. The mean voltage is −23.323 V and the variance is 0.03212 V. The voltage curve reflects the process of determining the optimal coupling position of the entire system in real time.

The simulated annealing algorithm found the best coupling position in a short time, and the coupling efficiency reached 51.4%. The piezoelectric ceramic driving circuit designed in this system has high precision, and the piezoelectric ceramic fixing mode is reliable. This method can effectively improve the coupling efficiency of optical fiber.

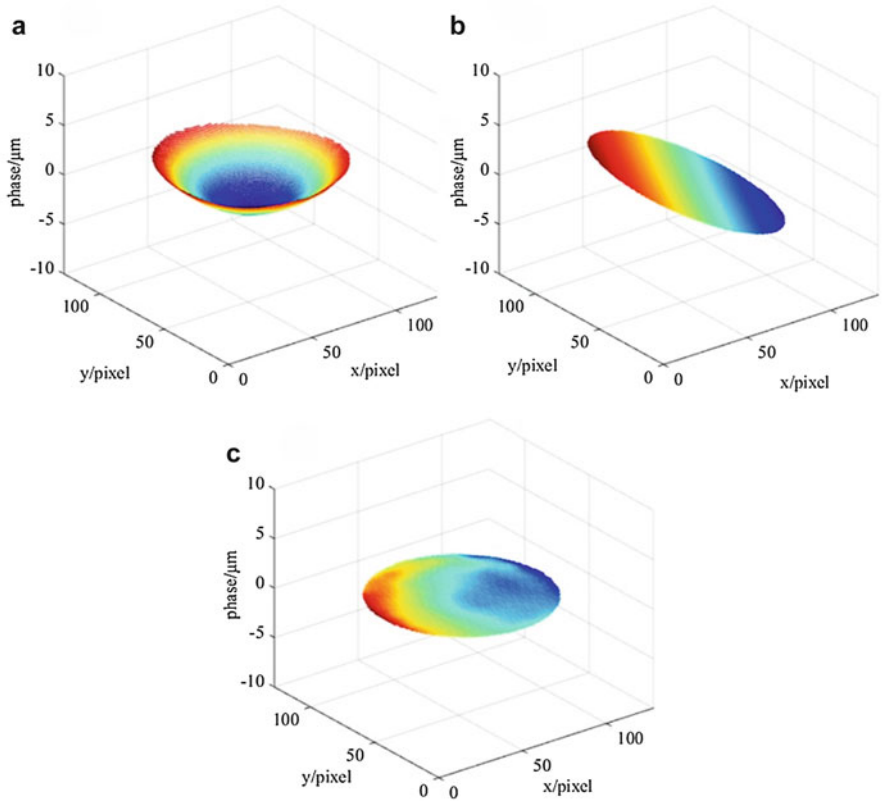


Fig. 23.26 Wavefront phase of an adaptive optical system based on a TM and DM combination. (**a**) TM correction; (**b**) DM69 correction; (**c**) simultaneous TM and DM69 correction

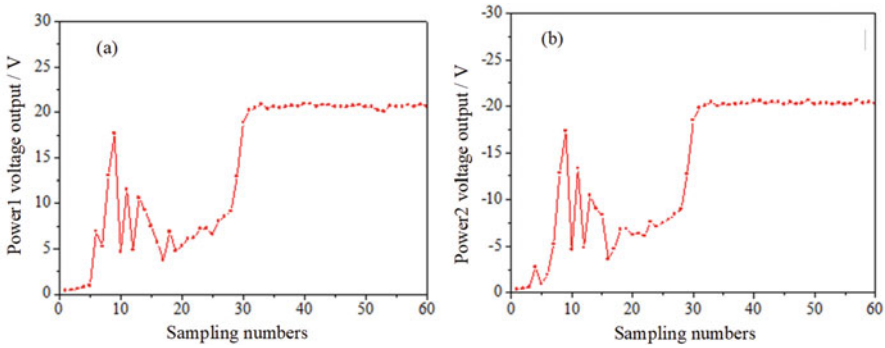


Fig. 23.27 Voltage variation when a simulated annealing algorithm is added [127]. (**a**) X axis; (**b**) Y axis

By applying two-dimensional piezoelectric ceramics to automatic fiber coupling alignment, Ke et al., Xi'an University of Technology, designed a five-degrees-of-freedom spatial light single-mode fiber coupling device, which adopted three different piezoelectric ceramic displacement modes, and designed the corresponding constant voltage DC power supply circuit and voltage amplification circuit, according to the working conditions of the three piezoelectric ceramics.

The SPGD algorithm was used to automatically align a spatial light and single-mode fiber coupling by considering the optical power in the coupled single-mode fiber as the system-performance evaluation function. Figure 23.28 shows the variation curve of the coupled optical power with the number of iterations under the introduction of different deviations.

Figure 23.28a shows that the initial coupled optical power is −12.37 dBm, when the axial deviation is introduced. After approximately 75 iterations of the algorithm, the optical power converges to about −1.15 dBm. As can be seen from Fig. 23.28b, the initial coupled optical power is −15.51 dBm when the radial deviation is introduced; the algorithm converges after about 90 iterations of the coupled optical power. After convergence, the coupled optical power fluctuates slightly around −1.21 dBm, indicating that the radial deviation is compensated.

As shown in Fig. 23.28c, when an angle deviation is introduced by adjusting the incident light direction, the initial coupled light power is −23.4 dBm. After

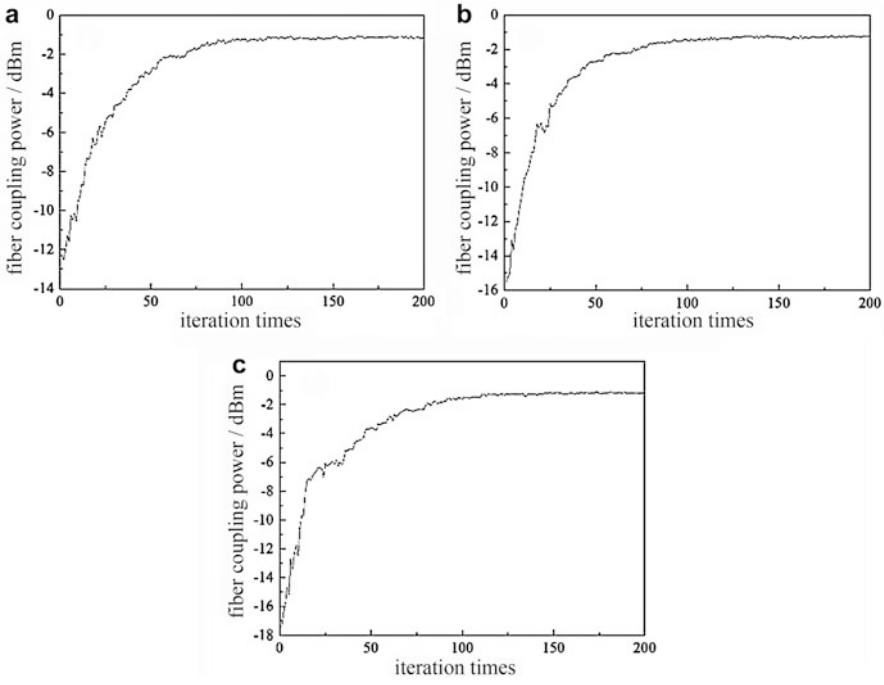


Fig. 23.28 Coupling optical power variation curve with the number of iterations under different deviations [128]. (**a**) axial; (**b**) radial; (**c**) angle

approximately 100 iterations of the algorithm, the coupled light power gradually converges and finally fluctuates around −1.19 dBm, indicating that the angle deviation is compensated [128].

Mode Conversion

In an optical wireless communication system, the spatial light mode is affected by atmospheric turbulence, and the beam mode is degraded, which leads to a mismatch between the mode field of the single-mode fiber and that coupled by the receiver. This reduces the efficiency of coupling the space light and single mode fiber. Ke et al. from Xi'an University of Technology experimentally studied the conversion of free-space light from a high order mode to a low-order mode based on LC-SLMs [129, 130].

Figure 23.29 shows the conversion result of the high order mode to the low order mode. The LP02 mode distribution in the first column is compared with the ideal light intensity distribution, and the distribution and symmetry of the light field energy are consistent with the ideal case; however, high-order diffraction components are inevitably present. After the degraded mode is passed through LC-SLM2, the obtained LP01-mode light field distribution energy concentration is consistent with the ideal mode field distribution. In addition, the distribution of the light field becomes more concentrated with a decrease in the mode order.

It can be observed that after the spatial-mode conversion based on a LC-SLM combined with the SA algorithm, the obtained mode distribution of LP01 is consistent with the distribution under ideal conditions, and the energy is concentrated in the central position, which has an obvious conversion effect.

When the degraded mode obtained by preprocessing is LP22, the coupling power during a pure binary conversion is −23.53 dBm, and the coupling efficiency after SA-algorithm optimization increases from 55.92% to 66.21%. An LC-SLM is used to convert the mode, increase the content of the fundamental mode, and indirectly

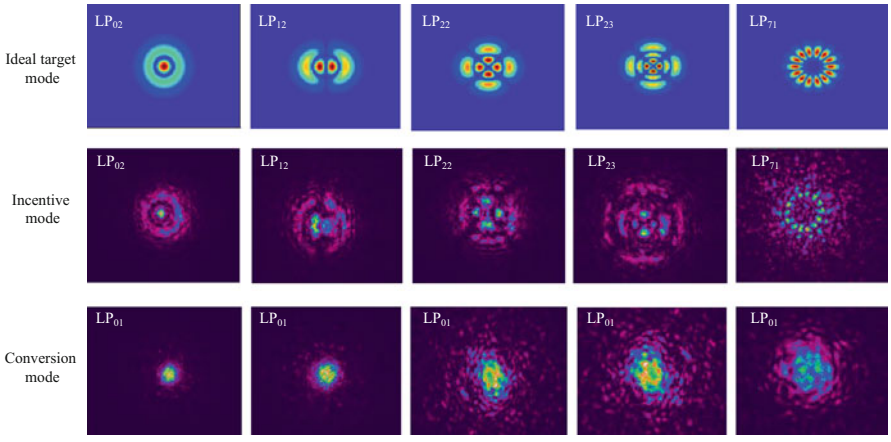


Fig. 23.29 Mode conversion results [130]

improve the coupling efficiency of the spatial optical single-mode fiber from the mode conversion. The coupling efficiency of the LP02 mode with the smallest order increases to 69.18%. The LP71 mode has the largest corresponding order, and the coupling efficiency increases from 49.12% to 60.29% of the binary conversion.

Adaptive Optical Fiber Coupling Experiment

To improve the efficiency of coupling space light to optical fiber, optical wireless communication using adaptive optics to suppress atmospheric turbulence was studied by combining adaptive and nonadaptive wavefront methods. Figure 23.30 shows the adaptive optical fiber coupling system. The adaptive optical wavefront system consists of a TM, DM69, wavefront sensor, and computer; the nonadaptive optical wavefront system consists of a splitting prism, coupling lens, single-mode fiber-optic power meter, and infrared camera.

Because the wavefront phase collected by the wavefront sensor is not in the same optical path as the fiber-coupled optical fiber end-face, the nonpublic optical path aberrations and fiber-coupled assembly errors generated by the lens are static errors. Their error magnitude is unrelated to changes in the external environment but is only related to the uniformity of the lens material, the device installation position, and the position of the fiber-coupled optical end-face.

Because of this error, the branch wavefront cannot be corrected after the closed loop of traditional adaptive optics. Two methods can eliminate this error: the blind optimization algorithm and the phase difference method. Compared with the phase difference method, the blind optimization method is more mature in practical engineering applications.

Figure 23.31 shows the coupled power curve and wavefront PV curve before and after error elimination by the SPGD algorithm. As shown in Fig. 23.31, after error elimination, the power decreases after the adaptive optical closed loop, whereas the coupled power increases after elimination.

Figure 23.32 shows the coupling power curve under adaptive optic open loop and closed loop conditions. As shown in Fig. 23.32, the coupling efficiency after

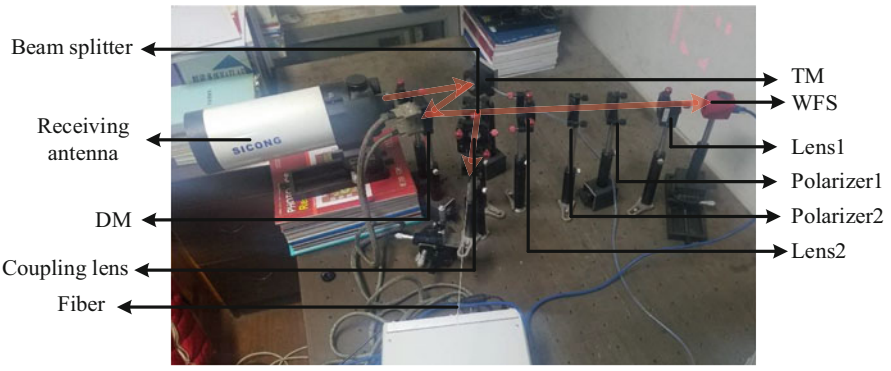


Fig. 23.30 Schematic diagram of an adaptive optical wavefront correction fiber coupling control system

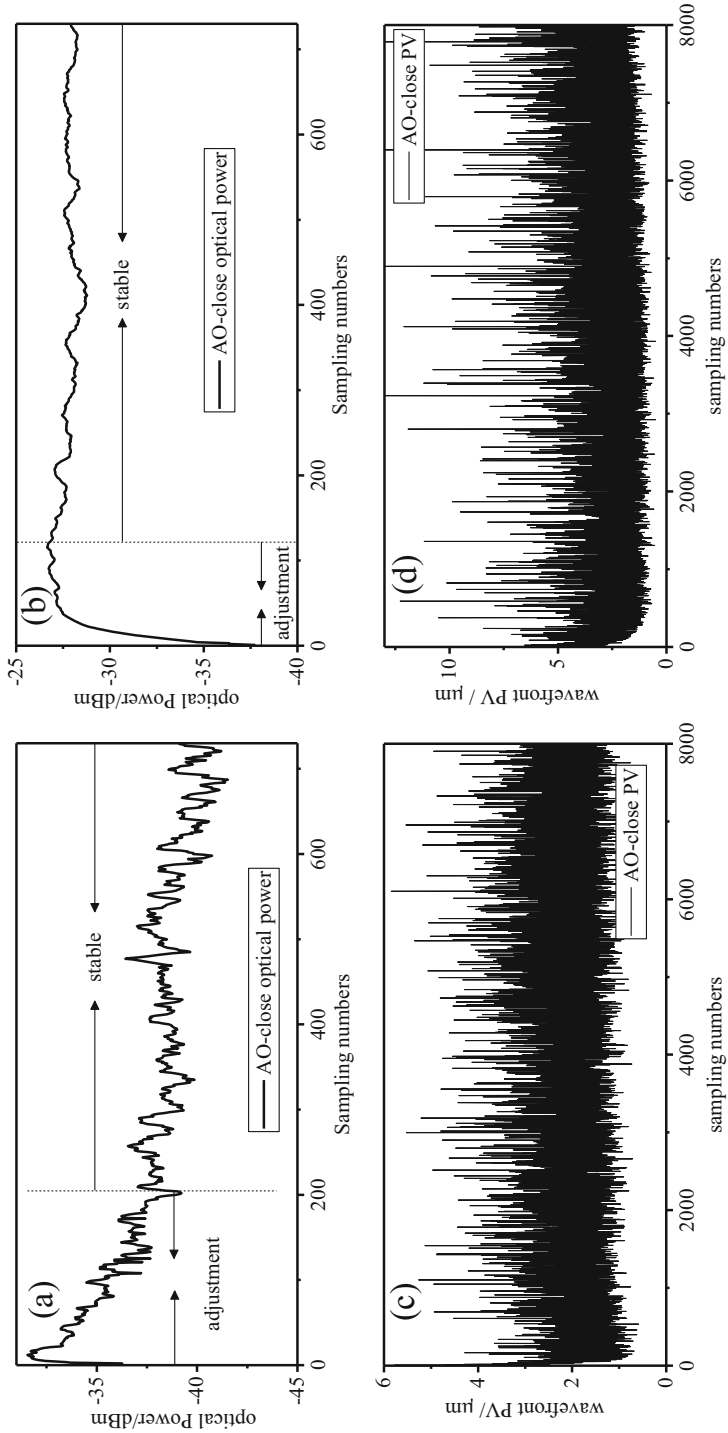


Fig. 23.31 Influence of a non-common optical path aberration on the adaptive optical closed loop coupling power and wavefront phase. (**a**) Uneliminated-aberration optical power; (**b**) eliminated-aberration optical power; (**c**) PV value of the aberration wavefront is not eliminated; (**d**) PV value of the aberration wavefront is eliminated

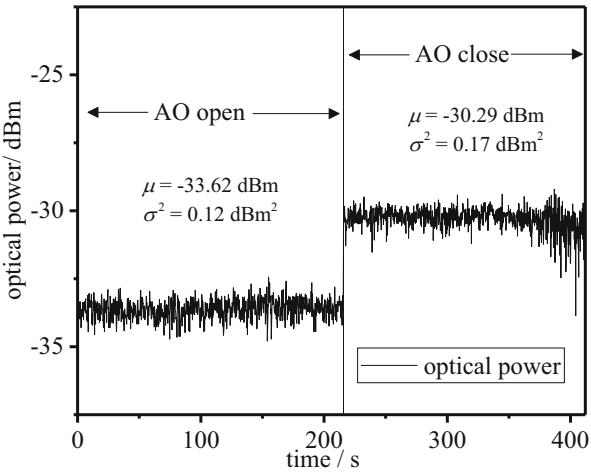


Fig. 23.32 Fiber coupling power variation curve before and after adaptive optical wavefront correction

correction is significantly higher than that before correction, indicating that adaptive optics can effectively improve the coupling efficiency of spatial light to fiber by correcting the distorted wavefront.

Observing the wavefront phase after eliminating the nonpublic optical path as the control object, Fig. 23.33 shows the calculation results of the coupling efficiency based on the Zernike coefficient of the measured wavefront at different distances. After the wavefront is corrected, the influence of wavefront distortion on the coupling efficiency increases from 10%, 5%, 10%, 3%, 5%, and 5% to 90%, 80%, 90%, 60%, 40%, and 10% when the optical wave is transmitted indoors, 600 m, 1 km, 5 km, 10 km, and 100 km, respectively. In addition, the effect of short-distance correction is better than that of long-distance correction. The numerical fluctuation of the coupling efficiency corrected at short distances is smaller than that at long distances.

For the coupling efficiency before and after correcting the 10-km wavefront, the correction had a numerical improvement in the coupling before the correction; however, the fluctuation of the coupling efficiency was too large, which was not conducive to the stability of the communication system. However, for a 100-km wavefront coupling efficiency, there was no significant improvement after correction. This is because the light intensity flicker caused by 100 km of strong turbulence and light spots is completely broken, which directly leads to the discontinuous acquisition of Zernike-coefficient wavefront data. The distortion of the 100-km wavefront is discontinuous and too large to ensure the real-time correction of the wavefront.

For subsequent research on the application of adaptive optics technology in optical wireless communication systems, the research content should focus on the convergence degree of a spot in the focal plane of the coupling lens after correcting the distorted wavefront, so that it can be directly coupled to the detector surface of the IM/DD system.

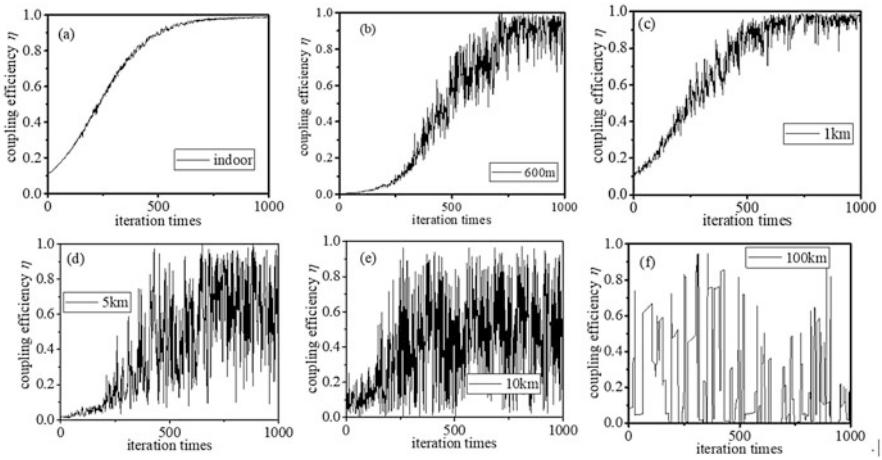


Fig. 23.33 Influence of wavefront correction on the coupling efficiency at different distances. (**a**) Indoors; (**b**) 600 m; (**c**) 1 km; (**d**) 5 km; (**e**) 10 km; (**f**) 100 km

For a remote coherent detection system, the research focus should be on non-wavefront sensing correction. Mode and non-mode optimization algorithms can improve the coupling efficiency of spatial light to fiber while correcting the distorted wavefront. Moreover, the relative fluctuation variance of the optical coupling power of adaptive optics technology in atmospheric turbulence should be considered to ensure the stable communication performance of the system.

Summary

This chapter systematically summarized the application of adaptive optics in the optical wireless communication field at home and abroad, and introduced the work in related fields by Xi'an University of Technology. The research work included adaptive optical systems with wavefront measurement, adaptive optical systems without wavefront measurement, wavefront correction of liquid crystal spatial light modulators, wavefront correction using a combination of tilt and deformable mirrors, and wavefront correction of spatial optical fiber-coupled adaptive optics. Finally, future research directions were presented.

References

1. Rao Changhui, Jiang Wenhan. Effectiveness analysis of adaptive optical system for atmospheric turbulence compensation [J]. High Power Laser and Particle Beams, 1996, 8(4): 469–475.
2. Babcock H W. The Possibility of Compensating Astronomical Seeing [J]. Publications of the Astronomical Society of the Pacific, 1953, 65(386): 229–236.

3. Hu Lifa, Liu Chao, Wen Shen, et al. Adaptive optics technology in the research development of astronomical observation [J]. Chinese science: mechanical astronomy, physics, 2017, 47 (8) : 7–18.
4. Zhang Yudong, Jiang Wenhan, Shi Guohua, et al. Ophthalmology of adaptive optics application [J]. Chinese science: mechanical astronomy, physics, 2007, 37 (s1) : 68–74.
5. Jiang Wenhan, Yang Zeping, Guan Chunlin, et al. New progress of adaptive optics in inertial confinement fusion [J]. Chinese Journal of Lasers, 2009, 36(7): 1625–1634.
6. Yang Huizhen, Li, Xinyang Jiang Wenhan. Application progress of adaptive optics in atmospheric optical communication system [J]. Laser & Optoelectronics Progress, 2007, 44(1): 61–68.
7. Fried D L. Adaptive optics topical issue [J]. Journal of the Optical Society of America A, 1977, 67(67): 422–428.
8. Tyson R K. Adaptive Optics and Ground-to-space Laser Communications [J]. Applied Optics, 1996, 35(19): 3640–3646.
9. Tyson R K. Bit-error Rate for Free-space Adaptive Optics Laser Communications [J]. Journal of the Optical Society of America. A, 2002, 19(4): 753–758.
10. Tyson R K, Canning D E, Tharp J S. Measurement of the Bit-error Rate of an Adaptive Optics, Free-space Laser Communications System, part 1: Tip-tilt Configuration, Diagnostics, and Closed-loop Results [J]. Optical Engineering, 2005, 44(9): 096002.1–096002.6.
11. Wilks S C, Morris J R, Brase J M, et al. Modeling of Adaptive Optics-based Free-space Communications Systems [C]//Free-Space Laser Communication and Laser Imaging II. New York: Academic Press, 2002: 121–133.
12. Thompson C A, Kartz M W, Flath L M, et al. Free space optical communications utilizing MEMS adaptive optics correction [C]//Free-Space Laser Communication and Laser Imaging II. Seattle: WA, 2002: 129–138.
13. Weyrauch T, Vorontsov M A. Atmospheric Compensation with a Speckle Beacon in Strong Scintillation Conditions: Directed Energy and Laser Communication Applications [J]. Applied Optics, 2005, 44(30): 6388–6401.
14. Weyrauch T, Vorontsov M A. Free-space Laser Communications with Adaptive Optics: Atmospheric Compensation Experiments [J]. Journal of Optical and Fiber Communications Reports, 2004, 1(4): 355–379.
15. Wright M W, Roberts J E, Farr W H, et al. Improved Optical Communications Performance Combining Adaptive Optics and Pulse Position Modulation [J]. Optical Engineering, 2008, 47(1): 016003.1–016003.8.
16. Hemmati H, Chen Y J, Crossfield I. Telescope Wavefront Aberration Compensation with a Deformable Mirror in an Adaptive Optics System [C]//Free-Space Laser Communication Technologies XVIII. San Jose: Academic Press, 2006: 123–131.
17. Wilson K E, Wright M W, Lee S, et al. Adaptive Optics for Daytime Deep Space Laser Communications from Mars [C]//Digest of the LEOS Summer Topical Meeting. New York: Academic Press, 2005: 19–23.
18. Boroson D M, Biswas A, Edwards B L. MLCD: Overview of NASA's Mars laser communications demonstration system [C]//Free Space Laser Communication Technologies XVI. San Jose: Academic Press, 2004: 16–33.
19. Stewart J B, Murphy D V, Moores J D, et al. Comparing Sdaptive Optics Approaches for NASA LCRD Ground Station #2[C]//Proceedings of SPIE-The International Society for Optical Engineering. US: Melanie Ortiz, 2013: 86100M.1–86100M.11.
20. Wright M W, Morris J F, Kovalik J M, et al. Adaptive Optics Correction into Single Mode Fiber for a Low Earth Orbiting Space to Ground Optical Communication Link Using the OPALS Downlink [J]. Optics express, 2015, 23(26): 33705–33712.
21. Juarez J C, Young D W, Sluz J E, et al. Free-space Optical Channel Propagation Tests over a 147-km Link [C]//Atmospheric Propagation VIII. New York: SPIE, 2011: 80380B-1–80380B-10.

22. Heine F, Kämpfner H, Czichy R, et al. Optical Inter-Satellite Communication Operational [C]//Milcom 2010 Military Communications Conference. San Jose: Academic Press, 2010: 1583–1887.
23. Sodnik Z, Armengol J P, Czichy R H, et al. Adaptive Optics and ESA's Optical Ground Station [C]//Free-Space Laser Communications IX. San Diego: SPIE, 2009: 746406.
24. Berkefeld T, Soltau D, Czichy R, et al. Adaptive Optics for Satellite-to-ground Laser Communication at the 1m Telescope of the ESA Optical Ground Station, Tenerife, Spain [C]//Adaptive Optics Systems II. Marseille: SPIE, 2010: 77364C-1–77364C-8.
25. Gregory M, Troendle D, Muehlnikel G, et al. Three Years Coherent Space to Ground Links: Performance Results and Outlook for the Optical Ground Station Equipped with Adaptive Optics [C]//Free-space Laser Communication and Atmospheric Propagation XXV. Switzerland: SPIE, 2013: 746406-1–746406-10.
26. Arimoto Y, Toyoshima M, Toyoda M, et al. Preliminary Result on Laser Communication Experiment Using (ETS-VI) [C]//Free-Space Laser Communication Technologies VII. New York: Academic Press, 1995: 151–158.
27. Kudielka K H, Hayano Y, Klaus W, et al. Low-Order Adaptive Optics System for Free-Space Lasercom Design and Performance Analysis [C]//Proceedings of the 2nd International Workshop on Adaptive Optics for Industry and Medicine. Singapore: IEEE, 2015: 1–12.
28. Petit C, Vedrenne N, Michau V, et al. Adaptive Optics Results with SOTA [C]//IEEE International Conference on Space Optical Systems & Applications. New York: IEEE, 2016: 1–7.
29. Hashmi A J, Eftekhar A A, Adibi A, et al. Analysis of Adaptive Optics-based Telescope Arrays in a Deep-space Inter-planetary Optical Communications Link between Earth and Mars [J]. Optics Communications, 2014, 333(12): 120–133.
30. Pasupathi T, Selvi J A V, Samuel J N. Mitigation of low-order atmospheric turbulent effects using sensorless adaptive optics in terrestrial free space optical communication [C]//2016 International Conference on Emerging Trends in Engineering. Lille France: ICETSET, 2016: 1–6.
31. Carrizo C, Calvo R, Belmonte A. Proof of Concept for Adaptive Sequential Optimization of Free-space Communication Receivers [J]. Applied Optics, 2019, 58(20): 5397–5403.
32. Brady A, Rossler C, Leonhard N, et al. Validation of Pre-compensation under Point-ahead-angle in a 1 km Free-space Propagation Experiment [J]. Optics Express, 2019, 27(13): 17840–17850.
33. Baykal Y, Gokce M, Ata Y. Application of Adaptive Optics on Bit Rrror Rate of M-ary Pulse-position-modulated Oceanic Optical Wireless Communication Systems [J]. Laser Physics, 2020, 30(7): 076202-1–076202-8.
34. Toselli I, Gladysz S. Improving System Performance by Using Adaptive Optics and Aperture Averaging for Laser Communications in Oceanic Turbulence [J]. Optics Express, 2020, 28(12): 17347–17361.
35. Paillier L, Bidan R, Conan J, et al. Space-Ground Coherent Optical Links: Ground Receiver Performance with Adaptive Optics and Digital Phase-Locked Loop [J]. Journal of Lightwave Technology, 2020, 38(20): 5716–5731.
36. Ata Y, Korotkova O. Adaptive Optics Correction in Natural Turbulent Waters [J]. Journal of the Optical Society of America A, 2021, 38(4): 587–594.
37. Osborn J, Townson M, Farley O, et al. Adaptive Optics Pre-Compensated Laser Uplink to LEO and GEO [J]. Optics Express, 2021, 29(4): 6113–6132.
38. Leonhard N, Berlich R, Minardi S, et al. Real-time Adaptive Optics Testbed to Investigate Point-ahead Angle in Pre-compensation of Earth-to-GEO Optical Communication [J]. Optics Express, 2016, 24(12): 13157–13172.
39. Chu X, Qiao C, Feng X. Propagation of Gaussian-Schell Beam in Turbulent Atmosphere of Three-layer Altitude Model [J]. Applied Optics, 2011, 50(21): 3871–3878.
40. Jiang Wenhan, Zhang Yudong, Rao Changhui, et al. Research progress of adaptive optics in Institute of Opto-Electronics, Chinese Academy of Sciences [J]. Acta Optica Sinica, 2011, 31(09): 72–80.

41. Jiang Wenhan. Review on the development of adaptive optics [J]. Opto-electronic Engineering, 2018, 45(03): 7–24.
42. xzheng Ke, Pengfei Wu. Adaptive Optics Theory and Its Application in Optical Wireless Communication [M], Beijing: Springer spress.
43. Yang Huizhen, CAI Dongmei, Chen Bo. Wavefront sensing adaptive optics and its application in atmospheric optical communication [J]. Chinese Journal of Lasers, 2008, 35(5): 680–684.
44. Li Huan, Zhang Hongtao, Yin Fuchang. Adaptive compensation method for atmospheric turbulence in space laser communication system [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2008, 31(02): 1–3.
45. Xia Lijun, Li Xiaofeng. Wavefront correction experiment of atmospheric optical Communication based on adaptive optics [J]. Information and Electronic Engineering, 2010, 8(3): 331–335.
46. Wu Yunyun, Chen Erhu, Zhang Yu, et al. Experimental verification of improving performance of FSO by adaptive optics [J]. Optical Communication Technology, 2012, 36(4): 15–18.
47. Han Liqiang, Wang Qi, Xin Taike GUI, et al. Research on performance of free-space optical communication system based on adaptive optical compensation [J]. Journal of Applied Optics, 2010, 31(2): 301–304.
48. Wu Yunyun, Chen Erhu, Zhang Yu, et al. Simulation and Experimental analysis of adaptive optics in optical Communication [J]. Optical Communication Technology, 2012, 36(8): 52–55.
49. Li J, Zhang Z, Gao J, Sun J, et al. Bandwidth of Adaptive Optics System in Atmospheric Coherent Laser Communication [J]. Optics Communications, 2016, 359(01): 254–260.
50. Li Jiawei, Chen Weibiao. Research on bandwidth of adaptive optical system in coherent optical communication between satellites and Earth [J]. Chinese Journal of Lasers, 2016, 43(8): 248–253.
51. Liu W, Shi W, Wang B, et al. Free Space Optical Communication Performance Analysis with Focal Plane based Wavefront Measurement [J]. Optics Communications, 2013, 309(07): 212–223.
52. Wu Yunyun, Li Xinyang, Rao Changhui. Effect of atmospheric Turbulence aberration on Error Performance of coherent Optical Communication with Space Zero binary Phase shift Keying [J]. Acta Optica Sinica, 2013, 33(6): 66–72.
53. Liu W, Shi W, Yao K, et al. Fiber Coupling Efficiency Analysis of Free Space Optical Communication Systems with Holographic Modal Wave-front Sensor [J]. Optics and Laser Technology, 2014, 60(08): 116–127.
54. Li Z K, Cao J T, Zhao X H, et al. Combinational Deformable Mirror Adaptive Optics System for Atmospheric Compensation in Free Space Communication [J]. Optics Communications. 2014, 320(06): 162–168.
55. Liu W, Yao K, Huang D, et al. Performance Evaluation of Coherent Free Space Optical Communications with a Double-stage Fast-steering-mirror Adaptive Optics System Depending on the Greenwood Frequency [J]. Optics Express, 2016, 24(12): 13288–13302.
56. Cao J, Zhao X, Li Z, et al. Stochastic Parallel Gradient Descent Laser Beam Control Algorithm for Atmospheric Compensation in Free Space Optical Communication [J]. Optik-International Journal for Light and Electron Optics, 2014, 125(20): 6142–6147.
57. Liu C, Chen S, Li X, and Xian H. Performance Evaluation of Adaptive Optics for Atmospheric Coherent Laser Communications [J]. Optics Express, 2014, 22(13): 15554–15563.
58. Liu C, Chen M, Chen S, Xian H. Adaptive optics for the free-space coherent optical communications [J]. Optics Communications, 2016, 361(02): 21–28.
59. Huang J, Deng K, Liu C, et al. Effectiveness of Adaptive Optics System in Satellite-to-ground Coherent Optical Communication [J]. Optics Express, 2014, 22(13): 16000–16007.
60. Huang J, Mei H, Deng K, et al. Signal to Noise Ratio of Free Space Homodyne Coherent Optical Communication after Adaptive optics Compensation [J]. Optics Communications, 2015, 356(12): 574–577.
61. Li M, Cvijetic M. Coherent Free Space Optics Communications over the Maritime Atmosphere with Use of Adaptive Optics for Beam Wavefront Correction [J]. Applied optics, 2015, 54(6): 1453–1462.

62. Li M, Gao W, Cvijetic M. Slant-path Coherent Free Space Optical Communications over the Maritime and Terrestrial Atmospheres with the Use of Adaptive Optics for Beam Wavefront Correction [J]. Applied optics, 2017, 56(2): 284–297.
63. Chen M, Liu C, Xian H. Experimental Demonstration of Single-mode Fiber Coupling over Relatively Strong Turbulence with Adaptive Optics [J]. Applied Optics, 2015, 54(29): 8722–8276.
64. Zhao S M, Wang L, Zou L, et al. Both channel coding and wavefront correction on the turbulence mitigation of optical communications using orbital angular momentum multiplexing [J]. Optics Communications, 2016, 376(10): 92–98.
65. Chen M, Liu C, Rui D, et al. Experimental Results of 5-Gbps Free-space Coherent Optical Communications with Adaptive Optics [J], Optics Communications, 2018, 418(003): 115–121.
66. Rui Daoman, Liu Chao, Chen Mo, et al. Application of Adaptive Optics on the Satellite Laser Communication Ground Station [J]. Opto-Electronic Engineering, 2018, 45(3): 55–63.
67. Yang L, Yao K, Wang J, et al. Performance Analysis of 349-element Adaptive Optics Unit for a Coherent Free Space Optical Communication System [J]. Scientific Reports, 2019, 9(1): 1–11.
68. Chang H, Yin X, Cui X, et al. Performance analysis of adaptive optics with a phase retrieval algorithm in orbital-angular-momentum-based oceanic turbulence links [J]. Applied Optics, 2019, 58(22): 6085–6090.
69. Rui D, Liu C, Chen M, et al. Probability enhancement of fiber coupling efficiency under turbulence with adaptive optics compensation [J]. Optical Fiber Technology, 2020, 60(8): 102343-1–102343-6.
70. Chang H, Yin X, Yao H, et al. Low-Complexity Adaptive Optics Aided Orbital Angular Momentum Based Wireless Communications [J]. IEEE Transactions on Vehicular Technology, 2021, 1(1): 1–13.
71. Gu H, Liu M, Liu H, et al. An algorithm combining convolutional neural networks with SPGD for SLAO in FSOC [J]. Optics Communications, 2020, 475(15): 1–6.
72. Jiang L, Dai Z, Yu X, et al. Experimental Demonstration of a Single-Mode Fiber Coupling Over a 1 km Urban Path with Adaptive Optics [J]. Journal of Russian Laser Research, 2021, 42(3): 363–370.
73. Zhang S, R Wang, Wang Y, et al. Extending the detection and correction abilities of an adaptive optics system for free-space optical communication [J]. Optics Communications, 2021, 482(31): 1–18.
74. Chen M, Liu C, Rui D, et al. Performance verification of adaptive optics for satellite-to-ground coherent optical communications at large zenith angle [J]. Optics Express, 2018, 26(4): 4230–4242.
75. Liu W, Yao K, Chen L, et al. Performance analysis of coherent free space optical communications with sequential pyramid wavefront sensor [J]. Optics & Laser Technology, 2017, 100(09): 332–341.
76. Xizheng Ke, Wu Jiali, Yang Shangjun. Research progress and prospect of atmospheric turbulence for optical wireless communication [J]. Chinese Journal of Radio Science, 2021, 36(03): 323–339.
77. Xizheng Ke, Deng Lijun. Partial Coherent Light Transmission Theory in Optical Wireless Communication [M]. Beijing: Science Press, 2018.
78. Xizheng Ke, Yin Zhiyun. Coding Theory in Wireless Laser Communication System [M]. Beijing: Science Press, 2009.
79. Xizheng Ke, Ke Dong. Optical Wireless Communication [M}, Beijing Springer Press, 2022
80. Xizheng Ke, Chen Juan, Deng Lijun. Hollow Time Coding Theory for Wireless Optical MIMO Systems [M]. Beijing: Science Press, 2014.
81. Xizheng Ke. Principle and Application of Wireless optical orthogonal Frequency Division multiplexing [M]. Beijing: Science Press, 2019.

82. Xizheng Ke, Wu Jiali. Principle and Application of Wireless Optical Coherence Communication [M]. Beijing: Science Press, 2021.
83. Ke Xizheng. Ultraviolet self-organizing Network Theory [M]. Beijing: Science Press, 2011. (in Chinese)
84. Ke Xizheng, Wang Jiao. Generation, Transmission, Detection and Application of Vortex Beams [M]. Beijing: Science Press, 2018.
85. Ke Xizheng, Chen Jinni. Wireless Laser Communication heterodyne detection System and Detection Method [P]. China Patent: CN103051375A, 2013-04-19.
86. Tan Zhenkun, Ke Xizheng, Wang Jiao. Experimental research on wavefront correction of heterodyne detection system [J]. Chinese Journal of Scientific Instrument, 2018, 39(12): 121–131.
87. Wu J, Ke Xizheng. Development of adaptive optical correction and polarization control modules for 10-km free space coherent optical communications [J]. Journal of modern optics, 2019, 67(3): 189–195.
88. Ke Xizheng, Yang S, Wang J. Experimental study of free space coherent optical communication on 1km [C]//10th International Coference on Advanced Infocomm Technology (ICAIT). Stockholm Sweden: IEEE, 2018: 61–65.
89. Ke Xizheng, Chen X. Correcting Wavefront Distortion of Dual-Wavelength Beams Due to Atmospheric Turbulence with a Correction Coefficient [J]. Optics and Photonics Journal, 2020, 10(4): 64–77.
90. Ke Xizheng, Tan Z. Effect of angle-of-arrival fluctuation on heterodyne detection in slant atmospheric turbulence [J]. Applied Optics, 2018, 57(5): 1083–1090.
91. Tan Z, Ke Xizheng. Analysis of a heterodyne detection system affected by irradiance and phase fluctuations in slant atmospheric turbulence [J]. Applied Optics, 2018, 57(32): 9596–9603.
92. Boyer C, Michau V. Adaptive optics: interaction matrix measurements and real time control algorithms for the COME-ON project [C]//Amplitude and Intensity Spatial Interferometry. New York: Academic Press, 1990: 63–81.
93. Kasper M, Fedrigo E, Looze D P, et al. Fast calibration of high-order adaptive optics systems [J]. Journal of the Optical Society of America A, 2004, 21(6): 1004–1008.
94. Paschall R N, Anderson D J. Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements [J]. Applied Optics, 1993, 32(31): 6347–6358.
95. Zhang Danyu Adaptive optics wavefront distortion control and experimental research [D]. Xi'an: Xi'an University of Technology, 2023.
96. Polo A, Haber A, Pereira S F, et al. An innovative and efficient method to control the shape of push-pull membrane deformable mirror [J]. Optics Express, 2012, 20(25): 27922–27932.
97. Haber A, Polo A, Smith C S, et al. Iterative learning control of a membrane deformable mirror for optimal wavefront correction [J]. Applied Optics, 2013, 52(11): 2363–2373.
98. Liu Zhangwen, Li Zhengdong, Zhou Zhiqiang, et al Adaptive optical correction technology based on fuzzy control [J] Journal of Physics, 2016, 65 (01): 131–138.
99. Ke X, Zhang D. Fuzzy control algorithm for adaptive optical system [J] Applied Optics, 2019, 58(36): 9967–9975.
100. Li Youkuan, Chen Dongquan, Du Xiangwan Influence of atmospheric scintillation on adaptive optical correction [J] Intense Laser and Particle Beam, 2004, 16 (05): 545–550.
101. Yang Ke Research on wavefront sensing technology of computed light field imaging [D] Xi'an: Xi'an University of Technology, 2021.
102. Ke Xizheng, Yang Ke, Zhang Ying Experimental study on detecting wavefront distortion by CAPIS technology [J] Progress in Laser and Optoelectronics, 2019, 56 (12): 30–37.
103. Ke Xizheng, Zhang Yunfeng. Wavefront sensing adaptive system and the Method of Improving Convergence Speed by using the System [P]. China Patent: CN110365404A, 2019-10-26.

104. Li Mei. Experimental study of beam wavefront distortion correction by Eigenmode method [D]. Xi'an: Xi'an University of Technology, 2023.
105. Yu Ji, Dong Bing. Experimental Study on wavefront sensorless Adaptive Optical System Based on Deformable Mirror Eigenmode [J]. Acta Optica Sinica, 2015, 35(3): 294–301.
106. Ke X, Li M. Laser beam distorted wavefront correction based on deformable mirror eigenmodes [J]. Optical Engineering, 2019, 58(12): 126101.1–126101.9.
107. Ke Xizheng, Zhang Yunfeng, Zhang Ying, et al. Gpu acceleration of wavefront sensing adaptive wavefront correction system [J]. Laser & Optoelectronics Progress, 2019, 56(7): 96–104. (in Chinese)
108. Ma Huimin, Zhang Pengfei, Zhang Jinghui, et al. Stochastic parallel gradient descent algorithm for adaptive optics systems [J]. High Power Laser and Particle Beams, 2010, 22(6): 1206–1210.
109. Zommer S, Ribak E N, Lipson S G, et al. Simulated Annealing in Ocular Adaptive Optics [J]. Optics Letters, 2006, 31(7): 939–941.
110. Yang Huizhen, Li Xinyang, Jiang Wenhan. Comparison of several stochastic parallel optimization control algorithms for adaptive optics systems [J]. High Power Laser and Particle Beams, 2008, 20(01): 11–18.
111. Wu Jiali. Experimental research on Coherent Optical Communication System without wavefront detection [D]. Xi'an: Xi'an University of Technology, 2018.
112. Vorontsov M A. Decoupled stochastic parallel gradient descent optimization for adaptive optics: integrated approach for wave-front sensor information fusion [J]. Journal of the Optical Society of America A, 2002, 19(2): 356–368.
113. Wu Jiali, Ke Xizheng. Adaptive optical correction without wavefront sensor [J]. Laser & Optoelectronics Progress, 2018, 55(3): 030103-1–030103-7
114. Tan Zhenkun. Influence Factors and Experimental Research on differential Detection Performance of Optical Wireless Communication [D]. Xi'an: Xi'an University of Technology, 2021.
115. Wang Xia-yao. Research on Adaptive optical correction technology of Vortex beam [D]. Xi'an: Xi'an University of Technology, 2018.
116. Ke Xizheng, Wang Xia-yao. Experimental study on correction of vortex wavefront distortion [J]. Acta Optica Sinica, 2018, 38(3): 197–203.
117. Cui Na Mei. Experimental study on Correction of wavefront Distortion of Vortex beam by phase difference method [D]. Xi'an: Xi'an University of Technology, 2023.
118. Ke X, Cui N. Experimental research on phase diversity method for correcting vortex beam distortion wavefront [J]. Applied Physics B, 2020, 126(66): 1–11.
119. Ke Xizheng, Han Kona. Wavefront simulation and correction of liquid crystal spatial light modulator [J]. Laser & Optoelectronics Progress, 2019, 56(05): 170–175. (in Chinese)
120. Kong Yingxiu. Wavefront correction technology for spatially coherent Optical Communication based on LC-SLM [D]. Xi'an: Xi'an University of Technology, 2021.
121. Kong Yingxiu, Ke Xizheng, Yang Yuan. Study on the effect of atmospheric turbulence on space coherent optical communication [J]. Laser & Optoelectronics Progress, 2015, 52(8): 95–101.
122. Li Xinyang, Ling Ning, Chen Donghong, et al. Stability control of High speed slanting mirror in adaptive optics system [J]. High Power Laser and Particle Beams, 1999, 11(1): 35–40.
123. Li Youkuan, Chen Dongquan, Du Xiangwan. Full-field compensation simulation of adaptive optics for double deformable mirror [J]. High Power Laser and Particle Beams, 2000, 12(6): 665–669.
124. Ke Xizheng, Luo Jing, Lei Sichen. Method and implementation of spatial optical coupling automatic alignment [J]. Infrared and Laser Engineering, 2018, 47(1): 93–99.
125. Ke Xizheng, Chen Wei. Space optical Coupling detection Device Based on Fiber Array [P]. Chinese Patent: CN103209022A, 2013-07-19.
126. Luo Jing. Research on Space Optical Coupling Automatic Alignment Technology [D]. Xi'an: Xi'an University of Technology, 2018.

127. Ke X, Yin B. Experimental Research on Automatic Alignment and Control Algorithm of Spatial Light-Fiber Coupling [J]. International Journal of Optics, 2021, 2021(9): 1–9.
128. Ke Xizheng, Zhang Xutong. A Method for Improving the Coupling Efficiency of Single-mode Fiber Using Mode Conversion [P]. China Patent: CN110133803A, 2019-08-18.
129. Zhang Xutong. Research on Improving Coupling Efficiency of Single Mode Fiber by Mode Conversion Method [D]. Xi'an: Xi'an University of Technology, 2023.
130. Xizheng Ke, Xutong Zhang. Conversion of free-space optical path mode by spatial light modulator [J]. Optical Engineering, 2020, 59(1): 016109.1–016109.12.

Mode Methods in Adaptive Optics

24

Contents

Introduction 876
Research Progress at Home and Abroad 876
Mode Methods 886
  Zernike Mode 888
  Lukosz Mode 894
  Deformable-Mirror Eigenmode Method 895
  Legendre Mode 899
Wavefront-Distortion Correction Quality-Evaluation Method 902
Application of Mode Methods to Wavefront Correction 905
  Application of Zernike Mode 905
  Application of the Deformable-Mirror Eigenmode Method 906
Summary 906
References 911

Abstract

Adaptive optics can effectively correct the wavefront distortion caused by atmospheric turbulence in optical wireless communication systems. As an adaptive optical control algorithm, the mode method has a significant effect on improving the performance of adaptive optical systems. This chapter systematically describes the research progress of Zernike mode method, deformable mirror eigenmode method, Lukosz mode method and Legendre mode method in the field of adaptive optics at home and abroad, introduces the common mode method and correction principle in detail, and briefly introduces the quality evaluation method of wavefront distortion correction in adaptive optics system and the application of mode method in wavefront correction. The research emphasis of mode method is to reduce the error and improve the speed accuracy, which directly affects the ability of wavefront distortion correction, and then affects the performance of adaptive optics system.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_24

Keywords

Mode method · Adaptive optics · Zernike mode method · Wavefront distortion

Introduction

Adaptive optics can suppress the perturbation of laser signals in atmospheric turbulence transmission [1], with the aim of correcting the wavefront aberration caused by the disturbance of the beam. At present, the common adaptive optics control algorithms mainly include the mode method [2], the control algorithm based on the principle of geometrical optics, and the model-free optimization algorithm, etc. In this chapter, we mainly focus on the Zernike mode method in the mode method, and introduce other mode methods at the same time. The mode method can effectively suppress the wavefront phase aberration, improve the correction speed and correction accuracy of adaptive optics system, and then improve the quality of communication. Therefore, finding a method that can improve and optimize the mode method has great potential for improving the performance of adaptive optics systems.

Research Progress at Home and Abroad

In 1953, Babcock proposed the concept of adaptive optics. He also proposed the idea of using optical elements with variable surface shapes to generate corresponding conjugate wavefronts to compensate for dynamic wavefront errors. This idea laid the foundation for the development of adaptive optics [3].

In 1976, Noll discussed the general properties of Zernike polynomials, such as Fourier transforms, integral representations, and derivatives. The Zernike representation of the Kolmogorov turbulence spectrum is the most commonly used Zernike expression in optics [4].

In 1977, Fried et al. successfully introduced adaptive-optics technology into optical wireless communication systems. The adaptive-optics unit could correct wavefront distortions caused by atmospheric turbulence in real time [5].

In 1979, Jiang developed a 19-unit wavefront-correction system for the Shenguang high-power laser device, which was the first practical device of its kind in the world [6].

In 1990, Nicolas proposed an algorithm that used the Zernike expansion of a random weighted function to simulate an atmospheric-distorted wavefront and introduced and analyzed the algorithm. The results showed that the Zernike polynomial well simulated the wavefront distortion [7].

In 1990, Daniel et al. experimentally verified that the orthogonality of Zernike polynomials makes them very suitable for wavefront fitting and proposed a Zernike-fitting method using a set of orthogonal polynomials. Experimental results showed that this method could speed up the Zernike fitting [8].

In 1995, Hu et al. introduced the eigenmode method for deformable mirrors, described its construction and application principles, and proved its relationship with the direct-slope method. Finally, using a 37-element deformable mirror as an example, the eigenmodes were constructed [9].

In 1996, Glenn et al. provided an expression for non-Kolmogorov turbulence using Zernike polynomials. The results showed that increasing the power-rate index of the three-dimensional phase power-spectral density from two to four resulted in the inclusion of a higher proportion of wavefront energy in the tilt component. For power-law exponents greater than four, the variance of the Zernike coefficients was more dependent on the outer turbulence scale than on the power-law index. The longer the length of the outer turbulence scale, the greater the energy in the tilt term [10].

In 1997, Butts reconstructed the phase as a linear combination of Zernike polynomials and calculated the structural function of the subaperture slope. The results showed that the Zernike spectrum calculated from the data measured by the wavefront sensor was consistent with the standard Kolmogorov turbulence model [11].

In 1999, Marcial et al. studied the properties of Gaussian functions when fitting a wavefront, analyzed a specific set of Gaussian functions, and demonstrated their simple Fourier transform. Two simulated wavefronts were fitted simultaneously and compared with Zernike polynomials. The results showed that, using the Gaussian function, the accuracy of the fitting increased with the number of terms [12].

In 2000, Neil et al. introduced a practical implementation of a closed-loop adaptive-optics system based on a novel wavefront sensor, where a light-spot pattern was generated on the detector plane. The experimental results showed that as the system aberrations were corrected, the Strehl ratio rapidly increased, and the light-spot energy was more concentrated [13].

In 2000, Zhang et al. used circular-domain and annular-domain Zernike polynomials to perform mode-method wavefront-reconstruction experiments on annular laser beams. The center occlusion ratio of the annular laser beam used in the experiment was 1:2. The results showed that if the wavefront-phase distortion of the annular laser beam was mainly composed of low spatial-frequency components, the wavefront could be directly reconstructed using a circular-domain Zernike polynomial. If the wavefront-phase distortion of the annular laser beam contained more high-spatial-frequency components, the annular-domain Zernike polynomial should be used for wavefront reconstruction [14].

In 2001, Hofer et al. used a real-time Shack–Hartmann wavefront sensor, a low-coherence source, and a scanning system to calculate the wavefront-phase root-mean-square value, Seidel aberrations, and average time power spectrum of 32 Zernike coefficients. The results showed that, except for defocus, human-eye aberrations fluctuated, and an adaptive-optics system with a closed-loop bandwidth of 1–2 Hz could accurately correct these aberrations [15].

In 2002, Chang et al. proposed a method for estimating Seidel-aberration coefficients using a two-dimensional discrete-wavelet transform. This method was used to analyze the wavefront of an optical system; it not only obtained more accurate

Seidel-aberration coefficients but also expedited the calculation. The experiment simultaneously fit three simulated wavefronts, and showed the simulation results of spherical aberrations, coma aberrations, astigmatism, and defocus [16].

In 2003, Mahajan expanded a set of Zernike polynomials with aberrations orthogonal to the system aperture; the variance was given by the sum of squares of the aberration coefficients. The experimental results showed that the Zernike polynomial was orthogonal to the circular pupil. For small aberrations, the Strehl ratio of the imaging system depended on the variance of the aberration [17].

In 2004, Upton et al. (2004) used the Gram–Schmidt orthogonalization technique to generate orthogonal hexagonal Zernike polynomials from circular-domain Zernike polynomials and provided the results of the first 17 sixth-order Zernike polynomials. Experiments confirmed that the Gram–Schmidt orthogonalization technique could be extended to any aperture shape [18].

In 2005, Kawashima et al. used a deformable mirror (DM) to compensate for aberrations in the human eye. (The working principle of a DM is to apply a control voltage to each actuator to change the shape of the deformable mirror, to correct a wavefront aberration.) The model-eye aberration measured by the Shack–Hartmann wavefront sensor was extended to a Zernike polynomial. The driving voltage of the deformable mirror was determined, and the human-eye aberration was canceled by using an appropriate voltage for each Zernike polynomial. The results showed that this method could reduce the root-mean-square error of corrected model human-eye aberrations to less than 0.1 μm [19].

In 2005, Seifert et al. studied the wavefront reconstruction of adaptive Shack–Hartmann sensors. The Zernike wavefront reconstruction was studied according to the fitting accuracy, reconstruction time, and noise influence. The experimental results showed that Zernike reconstruction was more suitable for fitting simple wavefronts [20].

In 2005, Yang et al. studied the reasons, conditions, and degree of mode coupling and mode confusion when reconstructing a wavefront using the Zernike mode method, and determined the basic principle of the optimal-reconstruction order. The 19-element Hartmann restoration accuracy was evaluated, and it was concluded that mode coupling was the main factor affecting the accuracy [21].

In 2006, Luo et al. used the Zernike-mode method to experiment with distorted-wavefront reconstruction, and obtained the Zernike optimization order of the distorted-wavefront reconstruction. The relationship between different Zernike orders and the recovery wavefront error was investigated through numerical simulations. The results showed that with an increase in the Zernike order, both the peak–valley (PV) and root-mean-square (RMS) values decreased rapidly. However, when the order exceeded 20, the change was relatively small, and the large number of calculations produced by the increase in the order also resulted in calculation errors [22].

In 2006, Booth proposed an adaptive-optics system that did not use wavefront sensors, but instead represented wavefront-phase distortions with appropriate mathematical models, such as Zernike polynomials. The research results showed that this method was helpful for optimizing control algorithms and was illustrated by the

application of direct-search and hill-climbing algorithms. Moreover, an efficient scheme using a maximization calculation was proposed, which allowed the measurement of N-order Zernike modes using only $N + 1$ intensity measurements [23].

In 2006, Fernandez et al. proposed a new deformable mirror using 52 independent magnetic actuators and studied in detail the ability of the deformable mirror to reproduce different surfaces, particularly Zernike polynomials to the fifth order. A study of the deformable-mirror influence function showed that the deformable mirror had a significant linear response to the applied voltage, and the correction device exhibited high fidelity when generating the surface. The range of Zernike polynomials completely covered common aberrations in the human eye, and even highly distorted eyes had almost perfect aberration correction [24].

In 2007, Martin et al. proposed a wavefront-sensorless adaptive-optics (WSAO) system based on the Lukosz mode (a new form developed after the derivative orthogonalization calculation of its Zernike polynomial) and introduced a scheme that could effectively measure wavefront aberrations represented by a series of modes. The scheme used the expansion of a Lukosz polynomial, based on the representation of aberrations. The experimental results showed that wavefront aberrations of the N-order aberration mode could be corrected with at least $N + 1$ photodetector measurements [25, 26].

In 2007, Rocha et al. measured changes in vision caused by individual Zernike aberrations using different RMS values. Experimental results showed that an electromagnetic adaptive-optics vision emulator could effectively correct and generate wavefront aberrations of up to four orders. Compared to the best eyeglass correction, custom wavefront correction significantly improved vision; and symmetric aberrations (such as defocus and spherical aberrations) were more detrimental to visual performance [27].

In 2007, Durán et al. introduced a scheme using traditional spatial light modulators to compensate for common optical aberrations in the human eye. The experimental results were obtained by correcting aberration distortions in artificial eyes by introducing a single Zernike term or a phase plate with complex ocular-aberration patterns. This was a step toward compensating for eye aberrations using low-cost general-purpose spatial light modulators [28].

In 2007, Qi et al. introduced a Zernike-polynomial fitting method for interference wavefronts. Research results showed that this method obtained the optimal mode combination of the interference wavefront and effectively improved the wavefront-fitting accuracy [29].

In 2008, Yang et al. proposed an adaptive-optics control algorithm that combined a genetic algorithm and the Zernike mode. Experimental results showed that using this algorithm shortened the iteration time and improved the correction speed. The PV and RMS values were reduced from 3.38λ and 0.64λ before the aberration correction to 0.52λ and 0.06λ, respectively. The improvement effect was significant, and the far-field focal spots before and after the correction were more concentrated from dispersion, which was close to the ideal Airy spot [30, 31].

In 2009, Yang et al. conducted a theoretical analysis and discussed the improvement in convergence speed of the stochastic parallel-gradient descent (SPGD)

algorithm based on the Zernike mode. In addition, a 61-element deformable mirror was used as the wavefront corrector to establish a simulation model of an adaptive-optics system.

The results showed that if the system contained only low-order distortion, mode-method optimizations could obtain a better convergence-speed and correction effect. When many high-order components were present, the mode and region methods could be combined, which not only obtained a fast convergence of the mode-method correction, but also maintained the correction ability of the entire system. While the same correction effect was achieved, the convergence speed of the algorithm using the combined optimization of the pattern and region methods was significantly better than the convergence speed based on the region method [32].

In 2009, Wang proposed a wavefront-reconstruction method based on a Walsh series. Because using the Walsh series as a base produced a correlation between the wavefront expansion and the light-field intensity, the wavefront mode coefficient could be obtained by measuring the incident-ray intensity using only a single-mode fiber and phase-spatial light modulator. Because this method only used light-intensity data, the calculation speed was faster than that of the wavefront-reconstruction method using two-dimensional slope information [33].

In 2010, Liu et al. proposed a Zernike-mode gray release graph-reconstruction algorithm for nematic liquid-crystal adaptive-optical systems. The results showed that the PV and RMS values corrected by the mode method were reduced from 2.50λ and 0.610λ to 0.0334λ and 0.00845λ, respectively; they were much smaller than those of the slope method of 0.173λ and 0.048λ, respectively. Compared to the slope method, the Strehl ratio of the system corrected by the mode method was closer to the diffraction limit [34].

In 2010, Mahajan et al. proposed the theory of noncircular-domain wavefront analysis and applied it to a systematic comparison of wavefront analyses of ring-domain and circular-domain Zernike polynomials. The research results showed that unlike ring-domain Zernike coefficients, circular-domain Zernike coefficients usually changed with the number of polynomials in the expansion. Moreover, the wavefronts fitted with a certain number of circular-domain Zernike polynomials were approximately the same as those fitted with the corresponding ring-domain Zernike polynomials [35].

In 2010, Lee proposed an efficient method for estimating the orthogonal-aberration coefficients of noncircular pupils. The method was based on the fact that all the information needed to construct an orthogonal polynomial (via the Gram–Schmidt process) could be obtained numerically using a conventional least-squares process of fitting Zernike polynomials to wavefront data. This enabled the method to obtain the required orthogonal-aberration coefficients using Zernike polynomial fitting. The method was well-suited for real-time wavefront analyses of dynamically changing circular pupils [36].

In 2010, Bueno et al. established a device consisting of a point-diffraction interferometer (PDI) and a Shack–Hartmann wavefront sensor to compare the two complementary techniques when applied to different phase plates. The results showed that under the same experimental and fitting conditions, the maximum

difference in RMS was 0.08 μm when all Zernike terms below order 6 were considered; when the lower-order term was excluded, the difference was reduced to 0.03 μm. The combination of these two techniques enabled the accurate measurement of complex phase distributions, while combining the robustness of the Shack–Hartmann sensor with the high resolution and dynamic range of a PDI [37].

In 2011, Chee et al. proposed a method to improve the accuracy of wavefront fitting. To obtain the coefficients of the Zernike polynomial, a set of discrete orthogonal polynomials needed to be constructed, and the coefficients were calculated by fitting the wavefront data and orthogonal polynomial using the least-squares method. The simulation results showed that, owing to the limitations of the least-squares method, there was a large deviation between the given coefficients and the calculated results [38].

In 2011, Rafael et al. proposed and evaluated complex Zernike polynomials (CZPs) to represent general wavefronts with non-uniform intensity (amplitude) in non-transmissive pupils. The results showed that the CZP provided a fully orthogonal basis capable of representing complex wavefronts, thereby providing a unified general framework. Complex wavefronts require more CZPs; however, they are particularly suitable for nonuniform light beams [39].

In 2011, Yu et al. proposed a simplified Gram–Schmidt method for solving Zernike polynomials; it had high fitting accuracy and was used to analyze the wavefront aberrations of interference fringes captured by interferometer systems. Experiments found that the stability of the Zernike coefficient changed with the Zernike term. When the Zernike term was less than 14, instability could be avoided [40].

In 2011, Feng et al. studied the influence of the order of Zernike polynomials on the fitting accuracy and conducted experiments. The experimental results showed that Zernike polynomials with appropriate orders should be selected to fit the wavefront to improve the fitting accuracy [41].

In 2011, Wang et al. studied the circular Zernike-polynomial pattern method to reconstruct the wavefront. The simulation results showed that the circular Zernike-polynomial mode method could ensure reconstruction accuracy. Moreover, wavefront-reconstruction experiments were performed based on Zernike polynomials with different numbers of terms. The results showed that with a change in the number of terms, the mode coefficients obtained by circular Zernike polynomials changed significantly [42].

In 2012, Chen et al. proposed a generalized ridge-estimation Zernike-mode wavefront-reconstruction algorithm for image restoration. Moreover, the generalized ridge-estimation point-spread function (PSF) and least-squares estimation PSF were compared. The results showed that the generalized ridge-estimation Zernike-mode method could effectively restore the optical wavefront. Using the generalized ridge-estimation PSF, the full width at half height of the restored image was 3.3 pixels with 48 iterations. Using the least-squares estimation PSF, the full width at half height of the restored image was 3.8 pixels with 65 iterations.

The restoration effect of the generalized ridge-estimation PSF was approximately 17% higher than that of the least-squares estimation PSF, which was closer to the

optical diffraction limit of the imaging system. It also reduced the number of iterations, which indirectly proved that the Zernike-mode reconstruction algorithm based on generalized ridge estimation had better PSF reconstruction accuracy [43].

In 2012, Booth et al. used a pattern-based adaptive-optics system to study the aberration correction of mouse embryos imaged in a microscope. The experimental results showed that the method had a fast correction speed and did not require multiple iterations; the real-time performance of the system was greatly improved [44].

In 2012, Hasan et al. studied Zernike polynomials for annular elliptical apertures in Cartesian and polar coordinates. These polynomials formed completely orthogonal bases on annular-elliptical apertures, representing balanced classic aberrations, like circular-field Zernike polynomials in a unit circle [45].

In 2013, Cacao et al. studied quaternionic Zernike spherical polynomials (QZSPs), presented their expressions, and summarized their basic properties. The results showed that the QZSP improved the original Zernike polynomials and was complete and orthogonal in the unit sphere [46].

In 2014, Yang et al. studied a model-based wavefront-detection adaptive-optics system and analyzed the correction effect and correction speed changes under different turbulence conditions. The research results showed that the system could obtain an ideal correction ability close to that of a wavefront-correction device while converging rapidly and had strong adaptability to different aberrations.

When an N-order mode aberration was corrected, the system only needed to measure $N + 1$ far-field light spots, which greatly reduced the number of measurements. This provided a feasible method for a wavefront-detection adaptive-optics system to be used for real-time aberration correction [47].

In 2014, Yan et al. designed an adaptive-optics system based on the deformable-mirror eigenmode method, and analyzed the influence of different orders of eigenmodes on the system-correction effect and closed-loop control stability under different thermal-halo intensities. Numerical-simulation results showed that the deformable-mirror eigenmode method could improve the stability of an AO-system closed-loop control and effectively suppress the occurrence of phase-compensation instability [48].

In the same year, Yan et al. designed an adaptive-optics system to correct the transient thermal-halo effect of collimated light beams. The experimental results showed that for a 69-element deformable mirror, when the order of the reconstructed mode was less than or equal to 37 orders, the greater the order of the reconstructed mode, the better the wavefront-correction effect. However, when the mode order was greater than or equal to 38 orders, the correction effect deteriorated rapidly [49].

In 2014, Yazdani et al. simulated two wavefronts with a typical arbitrary phase distribution. Numerical experiments verified an iterative method for simultaneously reconstructing two unknown interferometric wavefronts. An interferogram on two parallel planes perpendicular to the average propagation direction of these wavefronts was analyzed, and then the unknown wavefronts were calculated using the Zernike polynomial and the SPGD algorithm. The research results showed that

this method could effectively reconstruct two interferometric wavefronts simultaneously. It was verified experimentally to be a fast and accurate method [50].

In 2016, Chao et al. proposed a Gram–Schmidt orthogonalization method to handle the wavefront reconstruction of general aperture types. The method constructed a set of orthogonal basis functions and calculated the expansion coefficients to reproduce the original wavefront. Studies showed that the method was numerically stable, compared with the classic least-squares method, and exhibited superior performance for noncircular apertures. The method was robust to noise and did not cause truncation errors by omitting higher-order terms [51].

In 2016, Wang et al. proposed the use of a deformable-mirror eigenmode instead of the traditional Lukosz mode for a wavefront-sensor adaptive-optics system under point targets. The experimental results showed that after correction with the Lukosz mode, the remaining wavefront-error RMS value was 0.074λ, and after correction with the deformable-mirror eigenmode, the remaining wavefront-error RMS value was 0.042λ. The adaptive-optical correction method without a wavefront sensor, based on the deformable-mirror mode, could effectively correct aberrations with high correction accuracy and better than the traditional Lukosz mode [52].

In 2016, Zhang et al. used a Shack–Hartmann sensor and Zernike-mode reconstruction theory to conduct an experimental analysis of the light quality of a laser output beam. They accurately obtained the PV and RMS values of the wavefront and Zernike aberrations of each order. The far-field light-spot energy distribution was also obtained by further calculation so that the characteristics and quality of the light beam could be comprehensively analyzed. The experimental results showed that the aberration of the laser wavefront was concentrated in the first 10 stages of Zernike aberration [53].

In 2016, Dong et al. used a wavefront-sensorless adaptive-optics system based on the Lukosz mode to dynamically correct aberrations in infrared remote sensors. They used the spectral density of light-spot images at low spatial frequencies as a performance-index function, using 52-unit deformable mirrors for correction. The simulation results showed that the average image-sharpness value during the dynamic correction was 1.436×10^{-5}, and in the unoptimized correction, the average value of the image sharpness was 1.427×10^{-5}. They also verified that the mode-based wavefront sensing adaptive-optical system achieved a faster convergence than the traditional random parallel gradient-descent method [54].

In 2016, Li et al. proposed a wavefront-correction method using a focal-plane image in the field of optical wireless communication. The method divided the focal-plane image into 16 subregions, divided the first 18-order Zernike coefficients into 1736 subspaces, and constructed neural networks with only one hidden layer to successfully estimate the first 18-order Zernike coefficients [55].

In 2017, Wang et al. proposed a wavefront-distortion correction method based on the deformable-mirror eigenmode, and built an experimental system. The experimental results showed that the algorithm converged after three iterative corrections under different turbulence intensities, and the Strehl ratio of the system was higher than 0.8 [56].

In 2018, Liu et al. proposed using the difference Zernike-polynomial fitting (DZPF) method to reconstruct the wavefront from multiple phase differences. Numerical-simulation results showed that, compared with the discrete Fourier transform (DFT) method, the DZPF method maintained higher reconstruction accuracy in a wider shear-ratio range and had better noise immunity. Test experiments of the shear interferometer were also performed. In the wavefront-reconstruction process, the proposed DZPF method was used instead of the DFT method, and the measured RMS value of the shear interferometer was reduced from 0.0054λ to 0.0029λ [57].

In 2018, Wang et al. disclosed an adaptive-optics device based on the eigenmode control of a pyramid wavefront sensor. The eigenmode was used to obtain the transfer-function matrix of the adaptive-optics system, and mode filtering was used to improve the stability of the transfer-function matrix. Thus, it effectively improved the closed-loop stability of the adaptive-optics system. The correction effect of the system could be selected according to the actual situation [58].

In 2018, Ma et al. modified AlexNet [59] (a convolutional neural network for classification). First, the focal-plane and defocus-plane images under different atmospheric-turbulence parameters were generated by simulation. These images were used as the training dataset to train the modified AlexNet model, and the first 35-order Zernike coefficients were output [60].

In 2019, Sun et al. proposed a method for wavefront restoration using the Gram matrix equation and conducted experimental research. The experimental results showed that the method could recover the Zernike coefficient from the wavefront-layer information without pattern coupling. The method could be extended to the ring domain to obtain uncoupled ring-domain Zernike polynomial coefficients [61].

In 2019, Ke et al. used a wavefront-sensorless adaptive-optics system based on deformable-mirror (DM) eigenmodes to correct distorted far-field light spots and deduced a set of DM eigenmodes that conformed to the orthogonal relationship of derivatives. Combined with the principles of Fresnel diffraction and multiphase screen transmission, a system simulation model was established under different atmospheric-turbulence intensities. The results showed that this method could improve the signal-to-noise ratio of the system and effectively correct wavefront distortion caused by dynamic atmospheric turbulence [62].

In the same year, Ke et al. studied a computationally adaptive plenoptic imaging system (CAPIS) to detect optical wavefront distortions and carried out simulation analyses. The results showed that CAPIS technology could accurately detect low-order aberration wavefronts, and the RMS of the wavefront residuals were less than 0.1λ [63].

In 2020, Wu et al. designed an adaptive-optics system based on the deformable-mirror eigenmode and also conducted experimental research. The experimental results showed that, on the premise of ensuring the correction effect, regardless of the turbulence strength, the convergence speed of the system was improved to a certain extent. For the same wavefront aberration, the more deformable mirror units, the fewer required far-field light-spot measurements; the stronger the atmospheric-turbulence intensity, the more units required for deformable mirrors, and the fewer far-field light-spot measurements required [64].

In 2020, Liang et al. studied a wavefront-free adaptive-optics system based on the deformable-mirror eigenmode. Using the influence function of the deformable mirror, a set of deformable-mirror eigenmodes that conformed to the orthogonal derivative was deduced, and the eigenmodes were used as bases for decomposing the wavefront-aberration mode. The experimental results showed that the original wavefront PV value was 1.36 μm, the RMS value was 0.1995 μm, and the PV and RMS values after four corrections simultaneously reached optimal values, which were 0.514 μm and 0.088 μm, respectively. This method effectively corrected static aberrations, improved the energy concentration of far-field light spots, did not use multiple iterations, and had a fast convergence speed. Experiments required multiple measurements of far-field light-spot calculation mode coefficients and could only correct static aberrations well [65].

In 2020, Rodin et al. used convolutional neural networks to train and identify aberration types corresponding to a single Zernike polynomial, based on the intensity pattern of the point-spread function. They used a fast Fourier transform algorithm to simulate the point spread in the focal plane. When training neural networks, the learning coefficient and number of cycles of a dataset of a given size were empirically selected, and from a dataset of 17,000 point-spread function images, a set of 17 Zernike polynomials was obtained. For most types of aberrations, the average absolute prediction error was 0.012–0.017 [66].

In 2020, Wu et al. modified LeNet-5 [67] (a convolutional neural network for classification) and called the modified network a phase-diversity convolutional neural network (PD-CNN). They used a liquid-crystal spatial light modulator to generate the real phase, and experimentally acquired real images of the focal and defocal planes, outputting the first 13 Zernike coefficients. In the experimental verification, the trained neural networks were used to measure a wavefront with a PV value of 1.188λ and an RMS value of 0.226λ. The final recovered PV value was 0.078λ with an RMS value of 0.012λ [68].

In 2020, Hu et al. constructed neural networks based on U-Net [69] (a fully convolutional neural network) and ResNet [70] (a convolutional neural network), and named them SH-Net. The network inputs a light-spot array from the Hartmann sensor and directly outputs the phase map of the wavefront to be measured. It thereby realized an end-to-end mapping from the Hartmann sensor image to the phase map.

In the numerical simulation, the phase screen was generated by multiplying the Zernike-polynomial method and a Gaussian filter by a random-matrix method. In the test group, with the Zernike polynomial method and the random method simultaneously generating the phase screen, the wavefront RMS value restored by SH-Net was only 0.0168λ. In addition, the operational speed of the neural networks was considerable. The average operation time of SH-Net after training was 40.2 ms in actual use [71].

In 2021, Wang et al. proposed a wavefront-decoupling method for liquid-crystal deformable-mirror adaptive-optics systems based on the deformable-mirror eigenmode. The experimental results showed that the RMS of the correction residual after the eigenmode decoupling was 0.09 μm. The resolution of the 700 nm–1700 nm

wide band after decoupling correction increased from 1.1 times the diffraction limit to one times the diffraction limit, which was better than the 1.2 times the diffraction limit of the traditional Zernike mode [72].

In 2022, Cao et al. adopted a residual network as the backbone network and constructed a multi-scale residual mixed-attention network structure. They used the convolution operation to convert the light-intensity image into a feature map and propagated it backward; they extracted features through convolution kernels of different scales, and used the attention mechanism to improve the recognition rate of the network light-spot feature. Finally, the network-loss function was designed to obtain the Zernike coefficient that conformed to the actual wavefront aberration. The results showed that the residual attention network quickly and accurately reconstructed the turbulent phase, effectively correcting the distortion. This provides a guarantee for deep-learning practical applications in AO systems [73].

In 2022, Ning et al. proposed an adaptive-optics wavefront-correction method based on machine learning. This method used a turbulent-pool simulation to generate random disturbances and obtained sample data through wavefront detectors. This simplified the structure of the adaptive-optics system and improved the wavefront-correction accuracy and speed without wavefront sensing [74].

The mode-method application field extends from the optical wireless communication field to the medical field, with many studies and applications in the biomedicine field. In recent years, research on mode methods has mainly focused on error analysis, suppression, and elimination, as well as speed and accuracy. The speed and accuracy of the mode method directly affect the performance of the adaptive-optical system. Therefore, improving the solution speed and accuracy of the mode method will still be a research focus in the future.

For domestic development, most of the few related experiments at the beginning were numerical calculations and simulations. In recent years, new research methods have been proposed, and many new research directions have been added. However, this progress is insufficient. Most domestic development includes simulation analyses of algorithms. Experimental sites and equipments are limited. These are important factors that restrict development in China. The application of the mode method to adaptive optics still needs to be further researched, and improvements in the analysis are needed. Table 24.1 shows the research progress both at home and abroad.

Mode Methods

A mode method selects an appropriate base mode to describe the wavefront error and corrects the wavefront distortion by optimizing the base-mode coefficient [75, 76]. Currently, commonly used base modes include the deformable-mirror eigenmode [77], Zernike mode [4], Lukosz mode [77], and Legendre mode [78]. The base mode needs to satisfy the reciprocal orthogonal relationship [77].

Table 24.1 Research progress at home and abroad

Year	Researcher	Content	Result
1953	Babcock	Proposed the concept of adaptive optics.	–
1976	Noll	Defined the most commonly used Zernike expressions in optics.	The Zernike representation of the turbulent Kolmogorov spectrum was given.
1977	Fried	Introduced adaptive-optics technology into a optical wireless communication system.	The adaptive-optics unit corrected wavefront distortion caused by atmospheric turbulence in real time.
1979	Jiang	Developed a 19-unit wavefront correction system, 21-unit adaptive science system, and 2.16-m telescope infrared adaptive-optics observation system.	–
1990	Nicolas A R	Developed an algorithm for simulating atmospheric distorted wavefronts using a Zernike expansion.	Zernike polynomials simulated distorted wavefronts well.
1990	Daniel M H	Proposed a Zernike-fitting method using a set of orthogonal polynomials.	Expedited Zernike fitting.
1995	Hu	Introduced the deformable-mirror eigenmode method, and attempted to construct an eigenmode in practice.	–
1997	Butts	Reconstructed the phase as a linear combination of Zernike polynomials.	Zernike spectra calculated from wavefront sensor data were consistent with standard Kolmogorov turbulence models.
2000	Neil M A	Practically realized a closed-loop adaptive-optics system for a novel wavefront sensor.	SR increased rapidly with more concentrated light spots.
2005	Kawashima	Compensated for human-eye aberrations using deformable mirrors.	The RMS error of human-eye aberration correction was less than 0.1 mm.
2005	Yang	Created a Zernike-mode method to reconstruct wavefronts.	Mode coupling was the main factor affecting accuracy.
2006	Booth	Developed an adaptive-optics system without a wavefront sensor.	–
2007	Martin	Developed an adaptive-optics system without wavefront sensing based on the Lukosz mode.	The RMS phase error was reduced from 2.01 rad to 0.38 rad.
2008	Yang	Created an adaptive-optical control algorithm combining a genetic algorithm and Zernike mode.	PV and RMS values were reduced from 3.38λ and 0.64λ before correction to 0.52λ and 0.06λ, respectively, afterward.

(continued)

Table 24.1 (continued)

Year	Researcher	Content	Result
2009	Yang	Developed an SPGD algorithm based on a Zernike mode.	SR quickly increased to 0.8, requiring about 120 iterations.
2010	Liu	Created a Zernike-mode reconstruction algorithm for adaptive-optics systems.	Corrected PV and RMS values were much smaller than 0.173λ and 0.048λ.
2012	Chen	Created a generalized ridge-estimation Zernike-mode wavefront-reconstruction algorithm for image restoration.	Recovery was approximately 17% better than least-squares estimates while reducing the number of iterations.
2014	Yang	Designed a mode-based closed-loop control algorithm for adaptive-optics systems without wavefront detection.	Number of measurements was greatly reduced and correction speed was accelerated.
2017	Yu	Developed an adaptive-optics system without a wavefront sensor based on a deformable-mirror eigenmode method.	RMS phase error was reduced from 4.3 rad before correction to 0.26 rad after.
2016	Zhang	Experimentally analyzed the quality of a laser output light beam using a Zernike-mode reconstruction theory.	Laser wavefront aberrations were concentrated in the first 10 levels of Zernike polynomials.
2017	Wang	Developed wavefront-distortion correction based on deformable-mirror eigenmodes.	After correction, the SR was higher than 0.8.
2019	Ke	Created a wavefront-sensing adaptive-optics system based on deformable-mirror eigenmodes.	After six iterations, the system converged, and the SR was basically maintained at about 0.85.
2020	Wu	Created an adaptive-optics system based on a deformable-mirror intrinsic mode.	When $D/r_0 = 5$ and $D/r_0 = 10$, the system convergence speed increased by 2.39 times and 2.33 times, respectively.
2021	Rodin	Obtained the mean absolute error for each aberration of neural networks with a set of 17 Zernike functions using convolutional neural networks.	For most aberration types, the mean absolute error was 0.012–0.019.
2022	Ning	Developed an adaptive-optics wavefront-correction system based on machine learning.	Simplified the structure of the adaptive-optics system, and improved the wavefront-correction accuracy and speed without wavefront sensing.

Zernike Mode

Definition of a Zernike Polynomial

A Zernike polynomial is a set of completely orthogonal bases defined on a unit circle. At present, the most commonly used expression in optics is the Zernike polynomial defined by Noll [4], whose specific expression is

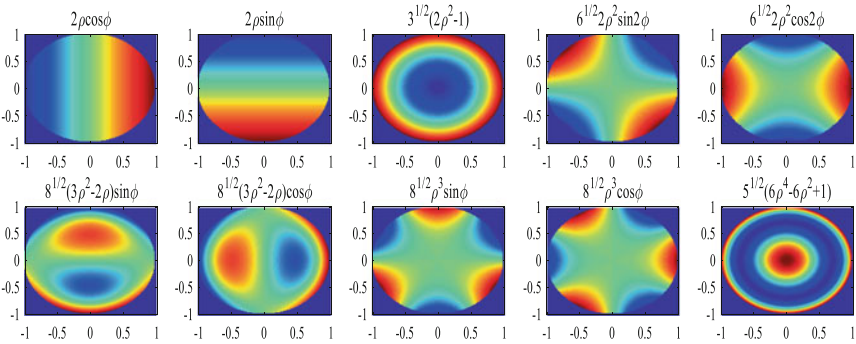


Fig. 24.1 Schematic diagram of the spatial distribution of Zernike modes of orders 2–10 [4]

$$Z_k(\rho,\theta) = \begin{cases} \sqrt{2(n+1)}R_n^m(\rho)\cos(m\theta) & k \text{ is an even number and } m \neq 0 \\ \sqrt{2(n+1)}R_n^m(\rho)\sin(m\theta) & k \text{ is an odd number and } m \neq 0 \\ \sqrt{n+1}R_n^0(\rho) & m = 0 \end{cases} . \tag{24.1}$$

In Eq. (24.1), $0 \leq \rho \leq 1$, $0 \leq \theta \leq 2\pi$, ρ and θ are the parameters of the polar coordinate system, n and m are the radial number and angular direction of the polynomial, respectively, and satisfy $n \geq 0$, $m \geq 0$, $m \leq n$; and n and m are even numbers. k is the serial number of the Zernike polynomial and $Z_k(\rho, \theta)$ is the expression of the Zernike polynomial. $R_n^m(\rho)$ is the radial function of the Zernike polynomial and satisfies

$$R_n^m(\rho) = \sum\nolimits_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k(n-k)!}{k!\left(\frac{n+m}{2}-k\right)!\left(\frac{n-m}{2}-k\right)!}\rho^{n-2k}. \tag{24.2}$$

Figure 24.1 shows a schematic diagram of the spatial distribution of Zernike modes of orders 2–10 obtained by MATLAB simulation.

Physical Meaning of Zernike Polynomials

The correspondence and physical meanings of the Zernike polynomials are listed in Table 24.2 [79]. From the correspondence between the Zernike polynomials and Seidel aberrations, the various aberration components and relative sizes contained in a distorted wavefront can be intuitively analyzed through the Zernike polynomial pattern and its coefficients.

Zernike Polynomials for Simulating Turbulent Phase Screens

Let the wavefront phase be $\varphi(x, y)$ and the coefficient of the Zernike polynomial of item i be a_i. A linear combination of several Zernike polynomials is used to represent an atmospheric-turbulence phase screen, and the expression is shown in Eq. (24.3):

Table 24.2 Correspondence between Zernike polynomials and low-order Seidel aberrations

Modulo ordinal number	Radial number	Angular number	Polar expression	Aberration
1	0	0	1	Translation
2	1	1	$2\rho\cos(\theta)$	Tilt in x direction
3	1	1	$2\rho\sin(\theta)$	Tilt in y direction
4	2	0	$\sqrt{3}(2\rho^2-1)$	Out of focus
5	2	2	$\sqrt{6}\rho^2\sin(2\theta)$	45° astigmatism
6	2	2	$\sqrt{6}\rho^2\cos(2\theta)$	0° or 90° astigmatism
7	3	1	$\sqrt{8}(3\rho^2-2\rho)\sin(\theta)$	Y-axis third-order coma
8	3	1	$\sqrt{8}(3\rho^2-2\rho)\cos(\theta)$	X-axis third-order coma

$$\varphi(x,y)=\sum_{i=1}^{n_Z}a_iZ_i(x,y). \tag{24.3}$$

In Eq. (24.3), a_i is the i-order coefficient of the adopted mode, Z_i is the adopted mode, n_Z is the Zernike polynomial of the former Z term, and $\varphi(x, y)$ is the wavefront phase. The covariance between the i-term and j-term Zernike polynomial coefficients is

$$\begin{aligned}\left\langle a_ia_j\right\rangle=(-1)^{(n_i+n_j-m_i-m_j)/2}\sqrt{(n_i+1)(n_j+1)}\delta_{m_imj}I_{n_in_j}\cdot\\ 2^{4-2\alpha}\pi B(\alpha)c(\alpha)(D_0/r_0)^{\alpha-2},\end{aligned} \tag{24.4}$$

where n_i, n_j, m_i, and m_j are the radial-frequency and angular-frequency numbers corresponding to Zernike polynomials of terms i and j, respectively. D_0 refers to the diameter of the telescope; r_0 is the atmospheric coherence length; and δ_{m_imj} is the expression of Kronecker logic symbols. $I_{n_in_j}$ B (α), c (α), δ_{m_imj} are respectively

$$\begin{aligned}I_{n_in_j}=&\frac{\Gamma\left[(n_i+n_j-\alpha+2)/2\right]}{\Gamma\left[(n_i-n_j+\alpha+2)/2\right]\Gamma\left[(n_j-n_i+\alpha+2)/2\right]}\cdot\\ &\frac{\Gamma(\alpha+1)}{\Gamma\left[(n_i+n_j+\alpha+4)/2\right]}\end{aligned} \tag{24.5}$$

$$B(\alpha)=\frac{\Gamma(\alpha/2)}{2^{2-\alpha}\pi\alpha\Gamma(-\alpha/2)}. \tag{24.6}$$

In Eq. (24.6), α is the absolute value of the refractive-index three-dimensional power-spectral density index term of non-Kolmogorov atmospheric turbulence, which satisfies $3<\alpha<4$. The Γ function is

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt, \tag{24.7}$$

where e is an exponential function and t is an intermediate variable of the integral.

$$c(\alpha) = 2\left[\frac{8}{\alpha - 2}\Gamma\left(\frac{2}{\alpha - 2}\right)\right]^{\frac{\alpha-2}{2}} \tag{24.8}$$

$$\delta_{m_i mj} = \left(m_i = m_j\right) \wedge [[(i + j) \in \text{even}] \vee (m_i = 0)] \tag{24.9}$$

In Eq. (24.9), even refers to an even number, $\delta_{m_i mj}$ is a Kronecker logic symbol, $\wedge$ is the "and" operation and $\vee$ is the "or" operation.

We perform a singular-value decomposition on the covariance matrix C_a. Because the covariance matrix has real symmetric positive determinism, its singular-value decomposition can be expressed as

$$C_a = XSX^T, \tag{24.10}$$

where S is a diagonal matrix and the transformation matrix X satisfies $X^{-1} = X^T$. In the simulation, a zero-mean random vector $\dot{d}$ was used with S as the covariance matrix. At this time, the following are obtained:

$$X\vec{d}\left(X\vec{d}\right)^T = X\vec{d}\vec{d}^T X^T = XSX^T = C_a. \tag{24.11}$$

$\dot{a} = X\dot{d}$ is the desired Zernike-polynomial coefficient vector.

The phase screen generated by this method has sufficient low-frequency components and relatively few high-frequency components. Although increasing the number of terms in the Zernike polynomial can solve this problem to a certain extent, it increases the calculation time.

Aberration of a Zernike Term Represents a Face Shape

An aberration means that, in an actual optical system, a single spherical lens or any combination of optical systems can only perfect the image of the paraxial-object point with a fine beam of light. The field of view and aperture increase and the concentricity of the imaging light beam is destroyed, resulting in various imaging defects. A Zernike polynomial is used to represent the surface shape of the optical element. The surface-shape error is

$$\phi(r, \varphi) = \sum_{i=0}^{N} A_i Z_i\left(\frac{r}{a, \varphi}\right), \tag{24.12}$$

where r and φ are the parameters of the surface-shape error, α is the radius of the optical element, Z_i is the Zernike polynomial, and A_i is the Zernike polynomial coefficient.

The definitions for the first 48 Zernike terms are as follows:

- Term 0 is the translation term, which represents the average optical-path difference.
- Terms 1 and 2 are tilt terms in the x and y directions, respectively.
- Term 3 is defocus.
- Terms 4 and 5 represent astigmatism and defocus, respectively.
- Terms 6 and 7 represent coma and tilt, respectively.
- Term 8 represents a spherical aberration.

- Terms 1–3 represent Gaussian or paraxial properties of the wavefront.
- Terms 4–8 represent three-order aberrations.
- Terms 9–17 are five-order aberrations.
- Terms 16–24 represent seven-order aberrations.
- Terms 25–35 represent nine-order aberrations.
- Terms 36–48 represent 11-order aberrations.

1. Tilt:

$$Z_1 = r\cos\frac{\varphi}{a}, Z_2 = r\sin\frac{\varphi}{a}. \tag{24.13}$$

2. Defocus:

$$Z_3 = \frac{2r^2}{a^2} - 1. \tag{24.14}$$

3. Astigmatism:

$$Z_4 = \frac{r^2}{a^2}, Z_5 = \frac{r^2\sin\varphi\cos\varphi}{a^2}. \tag{24.15}$$

4. Coma difference:

$$Z_6 = \frac{(3r^2-2)r\cos\varphi}{a^3}, Z_7 = \frac{(3r^2-2)r\sin\varphi}{a^3}. \tag{24.16}$$

5. Spherical aberration:

$$Z_8 = \frac{6r^4}{a^4} + \frac{6r^2}{a^2} - 1. \tag{24.17}$$

Principle of Zernike-Mode Wavefront-Distortion Correction

The mode method assumes that the distorted wavefront can be represented by a certain pattern [80]. First, the distortion information is converted into performance indicators, which are iteratively optimized using the optimization algorithm. The corrector is continuously deformed to obtain the optimal performance indicators for achieving the correction, which is the purpose of the distortion wavefront [81].

The evaluation function is obtained by calculating the integral of the image power-spectral density in the low-frequency region. According to the literature, the relationship between the image power-spectral density function and wavefront error is approximately [82]

$$g(M_1, M_2) \approx q_0(M_1, M_2) - q_1(M_1, M_2) \iint_p |\nabla \Phi|^2 ds, \tag{24.18}$$

where Φ is the wavefront error, $\nabla\Phi$ is obtained by differentiating Φ, q_0 and q_1 are constants related to the image structure, $\iint_p ds$ represents the integration of the pupil plane, and p refers to the pupil plane. M_1 and M_2 are normalized and represent the value range of the spatial frequency after processing. The wavefront error Φ is expanded through mode Z_i:

$$\Phi = \sum_{i=4}^{N} a_i Z_i. \tag{24.19}$$

In Eq. (24.19), a_i is the i-order coefficient of the adopted mode, Z_i is the adopted mode, and N represents the N-order coefficient. It is expanded to

$$\iint_p \nabla Z_i \cdot \nabla Z_j ds = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \tag{24.20}$$

where ∇Z_i and ∇Z_j are obtained by performing layer operations on Z_i and Z. The pupil surface is expressed by the mode coefficient as

$$G = g^{-1} = q_2 + q_3 \sum_{i=4}^{N} {a_i}^2, \tag{24.21}$$

where $q_2 = 1/q_0$ and $q_3 = q_1/q_0$.

The following system of equations can be derived when estimating the correction coefficient of the ith order mode:

$$\begin{cases} G_0 = q_2 + q_3 \sum_{k=i} a_k{}^2 + q_3 a_i{}^2 \\ G_+ = q_2 + q_3 \sum_{k=i} a_k{}^2 + q_3 (a_i + b_i)^2 \\ G_- = q_2 + q_3 \sum_{k=i} a_k{}^2 + q_3 (a_i - b_i)^2 \end{cases}. \tag{24.22}$$

In Eq. (24.22), G_0 is the performance index of the system in the initial state; G_+ and G_- are the performance indicators in the cases of statistical positive and negative bias, respectively; b_i is the offset; and a_i is the i-order coefficient of the adopted model. Solving this equation, we can obtain the correction of the i-order mode coefficient $a_{i,\ \mathrm{corr}}$:

$$a_{i,corr} = -a_i = -\frac{b_i (G_+ - G_-)}{2G_+ - 4G_0 + 2G_-}. \tag{24.23}$$

Lukosz Mode

The Lukosz polynomial is a new form, formed after the derivative-orthogonalization calculation of the Zernike polynomial [83, 84]. Its polar-coordinate expression is given by Eq. (24.24).

$$L_n^m(r,\theta) = B_n^m(r) \times \begin{cases} \cos(m\theta) & m \geq 0 \\ \sin(m\theta) & m < 0 \end{cases}, \tag{24.24}$$

where r and θ are polar-coordinate parameters, m is an integer, and n is a non-negative integer that satisfies $n - m = 2p$, $p = 0, 1, 2, \ldots$.

$B_n^m(r)$ can be expressed by Eq. (24.25):

$$B_n^m(r) = \begin{cases} \frac{1}{2\sqrt{n}} \left[R_n^0(r) - R_{n-2}^0(r) \right] & n \neq m = 0 \\ \frac{1}{2\sqrt{n}} \left[R_n^m(r) - R_{n-2}^0(r) \right] & n \neq m \neq 0 \\ \frac{1}{2\sqrt{n}} R_n^n(r) & n = m \neq 0 \\ 1 & n = m = 0 \end{cases}. \tag{24.25}$$

In Eq. (24.25), $R_n^m(r)$ is the amplitude of the Zernike polynomial, which can be expressed as Eq. (24.26):

$$R_n^m(r) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)! r^{n-2k}}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} \tag{24.26}$$

where both m and n satisfy the following relationship:

$$0 \le m \le n, n + m = 2p, p = 0, 1, 2 \ldots .$$

Figure 24.2 shows the wavefront diagrams of the Lukosz mode of the first 20 orders, and Table 24.3 is the expression of the first 20 orders.

Deformable-Mirror Eigenmode Method

Definition of the Deformable-Mirror Eigenmode Method

The deformable-mirror eigenmode is a mode with derivative orthogonality, obtained according to the influence function of the deformable mirror itself, which can

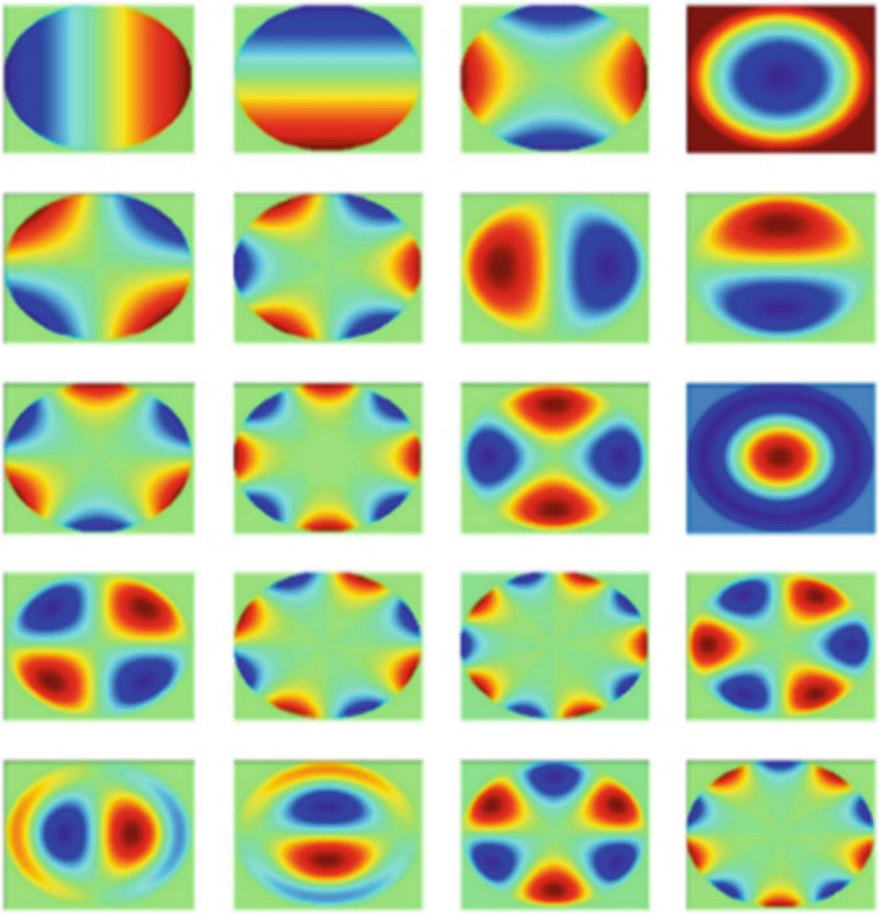


Fig. 24.2 Lukosz-mode wavefront diagrams of the first 20 orders [85]

Table 24.3 First 20 expressions of the Lukosz mode

Order	Lukosz-mode expression	Order	Lukosz-mode expression
1	1	11	$\frac{3}{2}r^4 - 2r^3 + \frac{3}{4}$
2	$r\cos\theta$	12	$\frac{\sqrt{2}}{2}r^2(r^2-1)\cos 2\theta$
3	$r\sin\theta$	13	$\frac{\sqrt{2}}{2}r^2(r^2-1)\sin 2\theta$
4	$\frac{\sqrt{2}}{2}(r^2-1)$	14	$\frac{1}{2}r^4\cos 4\theta$
5	$\frac{\sqrt{2}}{2}r^2\sin 2\theta$	17	$\frac{1}{2}r^4\sin 4\theta$
6	$\frac{\sqrt{2}}{2}r^2\cos 2\theta$	16	$\frac{\sqrt{10}}{2}r(2r^4-3r^2+1)\cos\theta$
7	$\frac{\sqrt{6}}{2}r(r^2-1)\sin\theta$	17	$\frac{\sqrt{10}}{2}r(2r^4-3r^2+1)\sin\theta$
8	$\frac{\sqrt{6}}{2}r(r^2-1)\cos\theta$	18	$\frac{\sqrt{10}}{2}r^3(r^2-1)\cos 3\theta$
9	$\frac{\sqrt{3}}{3}r^3\sin 3\theta$	19	$\frac{\sqrt{10}}{2}r^3(r^2-1)\sin 3\theta$
10	$\frac{\sqrt{3}}{3}r^3\cos 3\theta$	20	$\frac{\sqrt{5}}{5}r^5\cos 5\theta$

perfectly fit the deformable mirror shape. The specific derivation process is as follows [48].

The relationship between the surface shape ψ of the deformable mirror, the influence function ω of the deformable-mirror actuator, and the control voltage matrix C is

$$\psi = \omega \cdot C. \tag{24.27}$$

In Eq. (24.27), ψ is the surface shape of the deformable mirror; C is the control voltage matrix; ω is the influence function of the deformable-mirror actuator.

Then, the coupling matrix Γ between ω can be obtained, which is expressed as

$$\Gamma = \omega^T \cdot \omega, \tag{24.28}$$

In Eq. (24.28), ω^T is the transpose matrix and Γ is the coupling matrix between ω.

We decompose the singular value into

$$\Gamma = U \cdot S \cdot U^T, \tag{24.29}$$

In Eq. (24.29), U is the unitary matrix, U^T is the transpose matrix of U, and S is the diagonal matrix.

According to the characteristics of the unitary matrix, Eq. (24.27) can be expressed as

$$\psi = \omega \cdot U \cdot U^T \cdot C. \tag{24.30}$$

Let $M = \omega \cdot U$ and $C_M = U^T \cdot C$; then, Eq. (24.27) is

$$\psi = M \cdot C_M, \tag{24.31}$$

where M is the deformable-mirror eigenmode matrix and C_M is the coefficient matrix.

Correction Principle of the Deformable-Mirror Eigenmode Method

A distorted light spot is generated by the transmission of a Gaussian beam, the eigenmode of the deformable mirror is deduced as the base, and the system function is calculated to correct the wavefront aberration [86].

System Function

The power-spectral density $S_j(m)$ of the light-spot image at the receiving end is used, and the integral sum in the low-frequency part is used as a system function:

$$g(M_1, M_2) = \int_0^{2\pi} \int_{M_1}^{M_2} S_j(m) m dm d\xi, \tag{24.32}$$

where m is the normalized spatial frequency, $S_j(m)$ is the power-spectral density of the light-spot image, M_1 and M_2 are the value ranges of m [87], ξ is the spatial-frequency direction angle, and $g(M_1, M_2)$ is a system function. $\int_{M_1}^{M_2} S_j(m) m dm$ refers to the integral of the product of the power-spectral density of the light spot in the spatial-frequency range.

After the object passes through the incoherent-imaging optical system, the relationship between the object and the image can be expressed as

$$S_j(m) = |H(m)|^2 S_T(m), \tag{24.33}$$

where $H(m)$ is the optical-transfer function and $S_T(m)$ is the power-spectral density of the object. The relationship between the optical-transfer function (OTF) and the system aberration can be expressed as

$$|H(m)|^2 \approx |H_0(m)|^2 - \frac{1}{\pi} \iint_p (m \cdot \nabla \Phi)^2 ds, \tag{24.34}$$

where $\iint_p ds$ represents the integral of the pupil plane, p refers to the pupil plane, $H_0(m)$ represents the OTF in the ideal state, Φ removes the wavefront aberration after translation and tilt, and ∇ is the layer operator. The system function can then be expressed as

$$\begin{aligned} g(M_1, M_2) &= \int_0^{2\pi} \int_{M_1}^{M_2} S_j(m) m dm d\xi \\ &= \left\{ \int_{M_1}^{M_2} \left\{ |H_0(m)|^2 \int_0^{2\pi} S_T(m) d\xi - \frac{1}{\pi} \int_0^{2\pi} S_T(m) \left[\iint_p (m \cdot \nabla \Phi)^2 ds \right] d\xi \right\} m dm \end{aligned} \tag{24.35}$$

In Eq. (24.35), m is the spatial frequency after normalization, $m = (m \cos \xi, m \sin \xi)$, and $S_f(m)$ is the power-spectral density of the light-spot image. M_1 and M_2 are the value ranges of m, and ξ is the spatial-frequency direction angle. $H_0(m)$ represents the OTF in the ideal state and $S_T(m)$ is the power-spectral density of the object.

It can be seen that substituting $|\nabla\Phi|(\cos\chi, \sin\chi)$ into Eq. (24.29) yields

$$\begin{aligned}\iint_p (m \cdot \nabla\Phi)^2 ds &= \iint_p \{m|\nabla\Phi|[(\cos\xi \cdot \cos\chi) + (\sin\xi \cdot \sin\chi)]\}^2 ds, \\ &= \frac{m^2}{2} \iint_p (\nabla\Phi)^2 [1 + \cos(2\xi - 2\chi)] ds\end{aligned} \tag{24.36}$$

where χ is the spoke angle of $\nabla\Phi$, ξ is the spatial-frequency direction angle, Φ is the wavefront aberration after removing translation and tilt, $\nabla\Phi$ is obtained by differentiating Φ, $\iint_p ds$ represents the integral of the pupil plane, and p refers to the pupil plane.

The desired system-function expression can be obtained by calculation:

$$g(M_1, M_2) \approx q_0(M_1, M_2) - q_1(M_1, M_2) \iint_p |\nabla\Phi|^2 ds \tag{24.37}$$

$$q_0(M_1, M_2) = \pi \int_{M_1}^{M_2} |H_0(m)|^2 a_0(m) m dm \tag{24.38}$$

$$q_1(M_1, M_2) = \frac{\pi}{2} \int_{M_1}^{M_2} a_0(m) m^3 dm \tag{24.39}$$

Correction Coefficient

The system function g is represented by the mode coefficient C_{Mi} as

$$g = q_0 - q_1 \sum_{i=4}^{N} {C_{Mi}}^2, \tag{24.40}$$

where N is the mode order and C_{Mi} is the ith mode coefficient. When the aberration is large, Eq. (24.40) can be expressed as [88]

$$G = g^{-1} = q_2 + q_3 \sum_{i=4}^{N} {C_{Mi}}^2, \tag{24.41}$$

where $q_2 = 1/q_0$ and $q_3 = q_1/{q_0}^2$. Equation (24.41) is also suitable for small aberrations.

The relationship between the eigenmode coefficient of the ith deformable mirror and the system function is as follows:

$$\begin{cases} G = q_2 + q_3 \sum_{k=i} {C_{Mi}}^2 + q_3 {C_{Mi}}^2 \\ G_+ = q_2 + q_3 \sum_{k=i} {C_{Mi}}^2 + q_3 (C_{Mi} + b_i)^2 \\ G_- = q_2 + q_3 \sum_{k=i} {C_{Mi}}^2 + q_3 (C_{Mi} - b_i)^2 \end{cases}, \tag{24.42}$$

where C_{Mi} is the ith mode coefficient and G is the performance index of the system in the initial state.

Solving Eq. (24.42) yields the coefficient correction amount $C_{Mi,\mathrm{corr}}$ of the deformable-mirror eigenmode of the ith order:

$$C_{Mi,\mathrm{corr}} = -C_{Mi} = -\frac{b_i(G_+ - G_-)}{2G_+ - 4G_- + 2G_0}. \tag{24.43}$$

The above process is cycled N times to obtain the correction coefficient matrix $C_{Mi,\mathrm{corr}}$ of the N-order mode. If $C_M = U^T * C$ is known, the required correction voltage can be deduced.

Legendre Mode

Definition of the Legendre Polynomial

The Legendre polynomial [89] is the solution of the Legendre equation under the natural boundary condition: The solution $x = \pm 1$ is kept finite. The specific expression is as follows:

$$L_l(x) = \sum_{k=0}^{l/2} (-1)^k \frac{(2l-2k)!}{2^l k!(l-k)!(l-2k)!} x^{l-2k}, \tag{24.44}$$

where $l \geq 0$ and $[l/2]$ represents the largest integer not exceeding $l/2$. Using Eq. (24.44), the expression of the Legendre polynomial of the first seven orders can be obtained, as shown in Table 24.4.

Two-Dimensional Legendre Polynomials

Because the wavefront phase is distributed in a two-dimensional (2D) space, and the Legendre polynomial is defined in a one-dimensional space, when reconstructing the wavefront on the square domain, the Legendre polynomial needs to be extended to a two-dimensional space. The expansion method

involves multiplying two Legendre polynomials that are independent of each other in the x and y directions [90]:

$$L_{m,n}(x,y) = L_m(x) \times L_n(y), \tag{24.45}$$

where $L_m(x)$ and $L_n(y)$ are the mth and nth Legendre polynomials defined in the x and y directions, respectively.

A Fourier transform is performed on the two-dimensional Legendre polynomial, and the obtained frequency-domain analysis conclusion is the same as that of the Zernike polynomial. That is, as the two-dimensional Legendre order increases, the higher the spatial frequency it can represent. According to Eqs. (24.44) and (24.45), the first 10-order two-dimensional Legendre polynomial expressions can be obtained, as shown in Table 24.5. A schematic diagram of the spatial distribution of the 2D Legendre mode of orders 2–10 is shown in Fig. 24.3.

Unlike the Zernike polynomial, there is no direct connection between each term of the two-dimensional Legendre polynomial and various aberrations; therefore, it is impossible to analyze the distortion wavefront by the size of each coefficient, as can be done with the Zernike polynomial. Various aberration components and relative sizes exist.

Table 24.4 Legendre polynomials of the first seven orders

Serial number	Order	Cartesian coordinate expression
1	0	1
2	1	x
3	2	$(3x^2 - 1)/2$
4	3	$(5x^3 - 3x)/2$
5	4	$(35x^4 - 30x^2 + 3)/8$
6	5	$(63x^5 - 70x^3 + 17x)/8$
7	6	$(231x^6 - 317x^4 + 105x^2–5)/16$

Table 24.5 First 10-order two-dimensional Legendre polynomials

Serial number	$L_{m,n}$	m	n	Expression
1	$L_{0,0}$	0	0	1
2	$L_{1,0}$	1	0	x
3	$L_{0,1}$	0	1	y
4	$L_{2,0}$	2	0	$(3x^2 - 1)/2$
5	$L_{1,1}$	1	1	Xy
6	$L_{0,2}$	0	2	$(3y^2 - 1)/2$
7	$L_{3,0}$	3	0	$(5x^3 - 3x)/2$
8	$L_{2,1}$	2	1	$(3x^2 - 1)\ y/2$
9	$L1_{,2}$	1	2	$(3y^2 - 1)\ x/2$
10	$L_{0,3}$	0	3	$(5y^3 - 3y)/2$

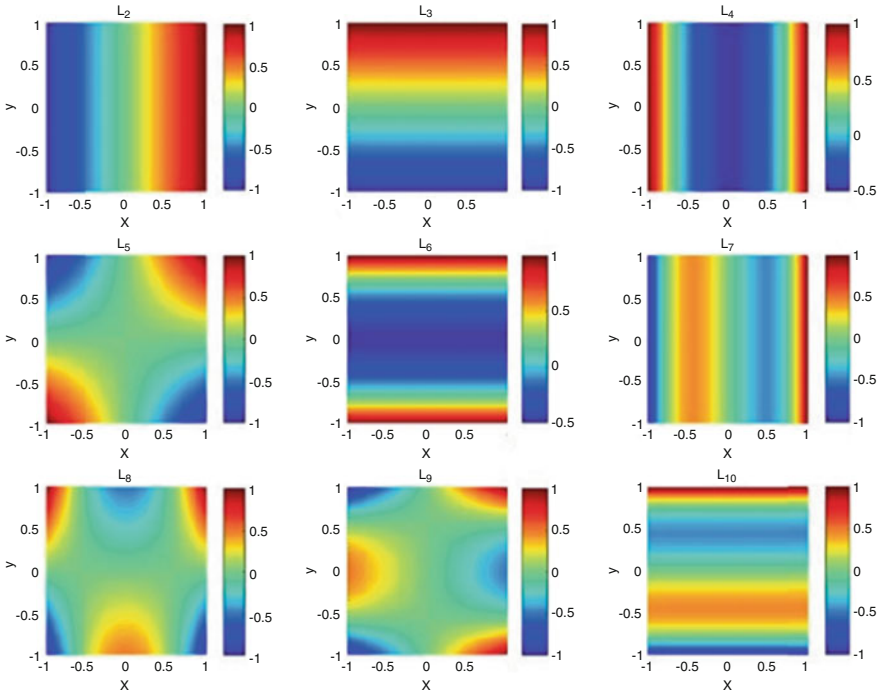


Fig. 24.3 First 10-order two-dimensional legendre polynomials [61]

Orthogonality of Two-Dimensional Legendre Polynomials

Legendre polynomials are orthogonal within their domain [84]. Therefore,

$$\int_{-1}^{+1} L_m(x)L_n(x)dx = \begin{cases} 0 & (n \neq m) \\ \dfrac{2}{2n+1} & (n = m) \end{cases}, \tag{24.46}$$

where m and n are polynomial orders. When the Legendre polynomial is extended to a two-dimensional space, if two two-dimensional Legendre polynomials are integrated in their domain, then, according to their extended definitions and the orthogonality of the Legendre polynomial, we can obtain

$$\int_{-1}^{1}\int_{-1}^{1} L_{m,n}(x,y)L_{k,l}(x,y)dxdy = \int_{-1}^{1} L_m(x)L_k(x)dx \int_{-1}^{1} L_n(y)L_l(y)dy, \tag{24.47}$$

where $L_{m,\,n}(x,\,y)$ and $L_{k,\,l}(x,\,y)$ are two-dimensional Legendre polynomials.

Wavefront-Distortion Correction Quality-Evaluation Method

An adaptive-optics system compensates and corrects the distortion of the optical wavefront. Its purpose is to correct the wavefront phase at each position in the optical receiving aperture to the same value. For point-target imaging, when the undistorted plane light wave is imaged through the optical system, the phase of each position in the optical receiving aperture is exactly the same, as shown in Fig. 24.4a. At this time, the light spot acquired on the imaging focal plane has the highest energy concentration and maximum light intensity, and the imaging effect at this time is the optical diffraction limit of the imaging system. The corresponding imaging light spot is called an Airy spot, as shown in Fig. 24.4b.

In contrast, if the light wave is distorted, the phase of each point in the optical receiving aperture is different, the energy concentration of the light spot obtained at this time decreases, the intensity of the center position of the light spot decreases, and the radius of the light spot increases. Therefore, the correction quality of the point light source imaging by the adaptive-optics system can be evaluated from three aspects: the phase difference of each point in the optical receiving aperture, the imaging light-spot radius, and the spot-center position intensity. The corresponding indicators are the wavefront-phase RMS error, optical diffraction-limit multiple, and spot Strehl ratio (SR).

The wavefront-phase RMS error describes the degree of fluctuation of the phase value of each point in the optical receiving aperture. The calculation method is as follows:

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[\varphi(x,y) - \langle\varphi(x,y)\rangle]^2}, \tag{24.48}$$

where $\varphi(x, y)$ is the wavefront phase and $\langle\varphi(x, y)\rangle$ is the average value of $\varphi(x, y)$.

The optical diffraction-limit multiple is the multiple of the light-spot radius obtained by the distorted incident-ray wave after passing through the imaging system, relative to the optical diffraction-limit light-spot radius. Let the intensity distribution of the imaging light spot be $I(x, y)$; then, the corresponding light-spot radius R can be calculated by the following expression:

$$\mathrm{R} = \frac{\iint\sqrt{(x-\bar{x})^2 + (y-\bar{y})^2}I(x,y)dxdy}{\iint I(x,y)dxdy}, \tag{24.49}$$

where $(\bar{x}, \bar{y})$ is the light-spot centroid coordinate. The smaller the optical diffraction-limit multiple, the higher the energy concentration of the light spot and the better the imaging quality. The Strehl ratio is defined as the ratio of the maximum intensity of

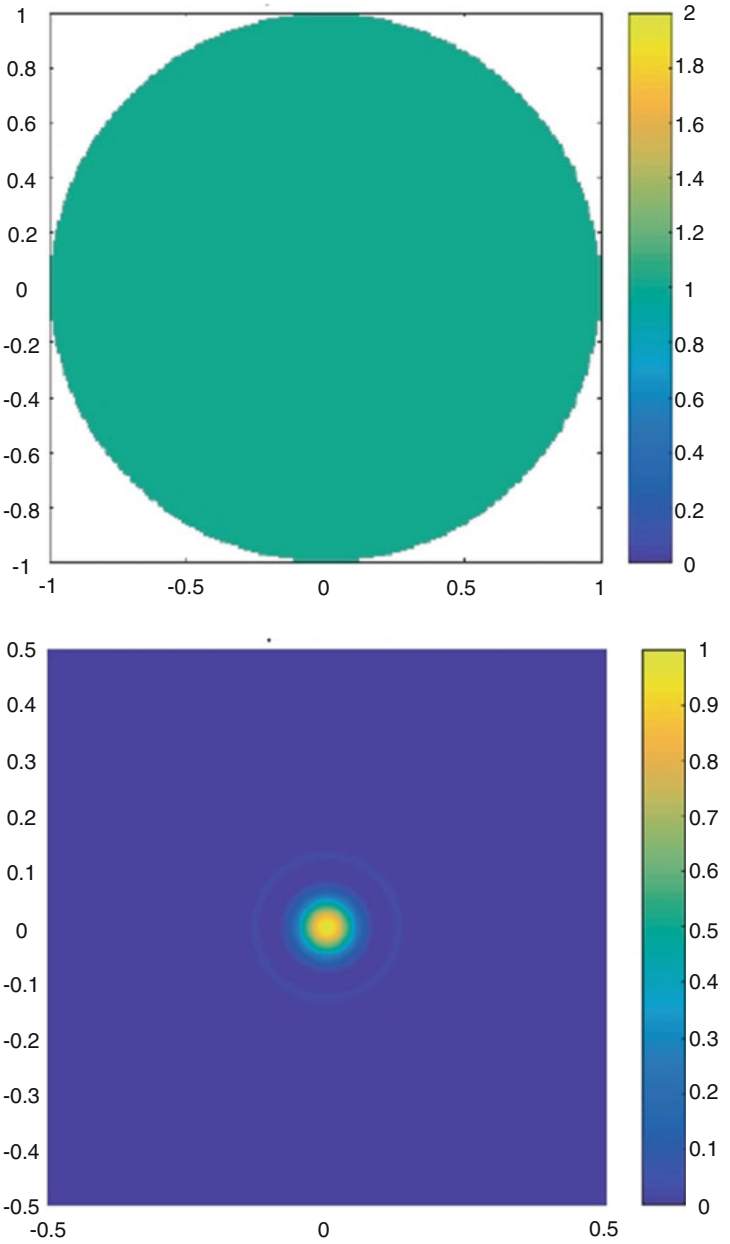


Fig. 24.4 Wavefront phase and imaging light spot without wavefront distortion [91]. (**a**) Wavefront-phase distribution. (**b**) Imaging light-spot intensity distribution

the actual imaging light spot to the maximum intensity of the ideal imaging light spot, as follows:

$$\mathrm{SR} = \frac{\max[I(x,y)]}{\max[I_0(x,y)]}. \tag{24.50}$$

In Eq. (24.50), x and y represent the coordinate values, $I(x, y)$ is the light-intensity distribution of the actual light spot with wavefront distortion, and $I_0(x, y)$ is the light-intensity distribution of the ideal imaging light spot without wavefront distortion. Max represents the maximum value. When the phase of the distorted wavefront $\varphi(\rho, \theta)$ is known, the Strehl ratio can also be calculated using the following expression:

$$\begin{aligned}\mathrm{SR} &= \frac{1}{\pi^2}\left|\int_0^{2\pi}\int_0^1 \exp[j\varphi(\rho,\theta)]\rho d\rho d\theta\right|^2 \\ &= \frac{1}{\pi^2}\left|\int_0^{2\pi}\int_0^1 \sum_{n=0}^{\infty}\frac{[j\varphi(\rho,\theta)]^n}{n!}\rho d\rho d\theta\right|^2,\end{aligned} \tag{24.51}$$

where $\varphi(\rho, \theta)$ is the distortion-wavefront phase. When $\varphi(\rho, \theta)$ is smaller, the higher-order terms of the third order and above in the Taylor expansion of Eq. (24.51) can be ignored so that

$$\begin{aligned}\mathrm{SR} &= \left|1 + j\langle\varphi(\rho,\theta)\rangle - \frac{1}{2}\langle\varphi^2(\rho,\theta)\rangle\right|^2 \\ &\approx 1 - \left[\langle\varphi^2(\rho,\theta)\rangle - \langle\varphi(\rho,\theta)\rangle^2\right] \approx \exp\left(-\mathrm{RMS}^2\right).\end{aligned} \tag{24.52}$$

When the Strehl ratio of the light spot is greater than 0.8, the quality of the imaging light spot is higher and the wavefront distortion is better corrected. If the imaging target is an extended target with a larger size, the intensity distribution of the light spot imaged by the point light source must be regarded as the point-spread function of the imaging system, i.e.,

$$f(x,y) = I(x,y) * h(x,y), \tag{24.53}$$

where $I(x, y)$ is the intensity distribution of the imaging target, $h(x, y)$ is the point-spread function of the imaging system, $f(x, y)$ is the imaging result of the target after passing through the optical system, and * is the convolution operator. The imaging result corresponding to the Airy spot is used as a reference image for the quality evaluation. Common metrics include the peak signal-to-noise ratio (PSNR) and structural similarity. The method for calculating the peak signal-to-noise ratio is as follows:

$$\mathrm{PSNR} = 10\log_{10}\left\{\frac{\left(2^k-1\right)^2}{\mathrm{MSE}[f(x,y),I(x,y)]}\right\}, \tag{24.54}$$

where k represents the number of bits in the image and MSE[$f(x, y)$, $I(x, y)$] is the mean-squared error between the imaging screen $f(x, y)$ and the reference image $I(x, y)$. Let the sizes of both images be $N_x \times N_y$, then, the mean-squared error is calculated as

$$\mathrm{MSE}[f(x,y),I(x,y)] = \frac{1}{N_x N_y}\sum_{x=0}^{N_x-1}\sum_{y=0}^{N_y-1}[f(x,y)-I(x,y)]^2. \tag{24.55}$$

The structural similarity is calculated as

$$\mathrm{SSIM}[f(x,y),I(x,y)] = \frac{\left(2\mu_f\mu_I+c_1\right)\left(2\sigma_{fI}+c_2\right)}{\left(\mu_f^2+\mu_I^2+c_1\right)\left(\sigma_f^2+\sigma_I^2+c_2\right)}, \tag{24.56}$$

where μ_f, μ_f^2, and μ_I, μ_I^2 are the mean and variance of the imaging picture and reference image, respectively, and σ_{fI} is the covariance between the two images. The values of parameters c_1 and c_2 are $(k_1L)^2$ and $(k_2L)^2$. $k_1 = 0.01$ and $k_2 = 0.03$, where L is the image pixel value and SSIM is the structural similarity.

Application of Mode Methods to Wavefront Correction

Application of Zernike Mode

Ke Xizheng et al. from Xi'an University of Technology studied the detection of optical wavefront distortion using a computationally adaptive plenoptic imaging system (CAPIS) [63]. As shown in Fig. 24.5, the RMS value of the single-order

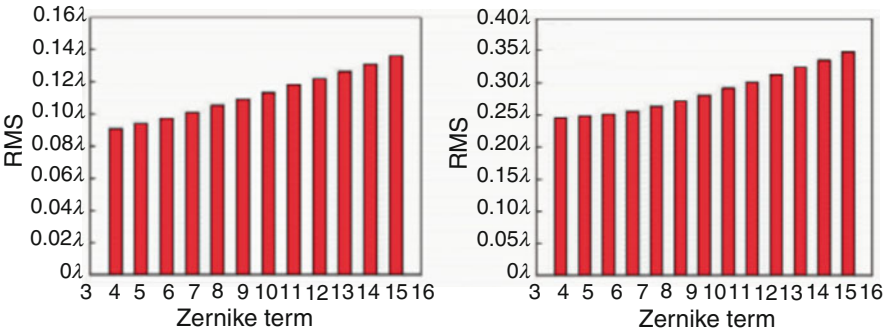


Fig. 24.5 RMS values of Zernike aberration detection results [63]. (**a**) Single-order aberration. (**b**) Combined aberration

aberration reconstruction is smaller and the accuracy is higher, whereas the RMS value of the combined aberrations is larger. It can be concluded that the effect of single-order aberration reconstruction is better than that of combined aberrations.

Ke and Zhang of Xi'an University of Technology designed an adaptive-optics system based on a double-fuzzy adaptive PID control algorithm [92]. Figures 24.6, 24.7, and 24.8 show the distribution curves of the Zernike coefficient values and wavefront PV values after wavefront correction with dual-fuzzy control under different parameters. Figure 24.6a shows the Zernike coefficient value and Fig. 24.6b shows the PV value. Experiments showed that the adaptive dual-fuzzy PID method can effectively correct the wavefront-distortion signal, improve the system response speed, and reduce the parameter-adjustment difficulty.

Application of the Deformable-Mirror Eigenmode Method

Ke and Li of Xi'an University of Technology built a wavefront-sensorless adaptive-optics (WSAO) experimental system based on the deformable-mirror eigenmode method to correct distorted light spots [62]. As shown in Fig. 24.9, after correction, the light spots become more convergent and the energy becomes more concentrated. This system can effectively correct the wavefront distortion caused by dynamic turbulence. As shown in Fig. 24.10, the system converges after six iterations, and the SR value is maintained at approximately 0.85.

Summary

This chapter systematically summarized the research progress on mode methods in the field of adaptive optics at home and abroad, pointed out their research directions, and provided a comparison of the research progress. It included an optical-system wavefront-distortion correction quality-evaluation method and the application of mode methods in wavefront-distortion correction.

It can be seen from the above discussion that the mode method is relatively mature, theoretically, in adaptive optics, and is developing in more directions. Astronomical observations and adaptive-optics system wavefront-distortion corrections are both application scenarios, as is medical human-eye aberration compensation.

In terms of research, methods based on the Zernike mode and deformable mirror eigenmode both exist. The research focus of mode methods is mainly on reducing errors and improving speed and accuracy, which are directly related to the size of the wavefront-distortion correction capability, which in turn affects the performance of the adaptive-optics system; thus, these aspects are still future research directions.

To further optimize the performance of a model-based adaptive optics system, the following adjustments can be made:

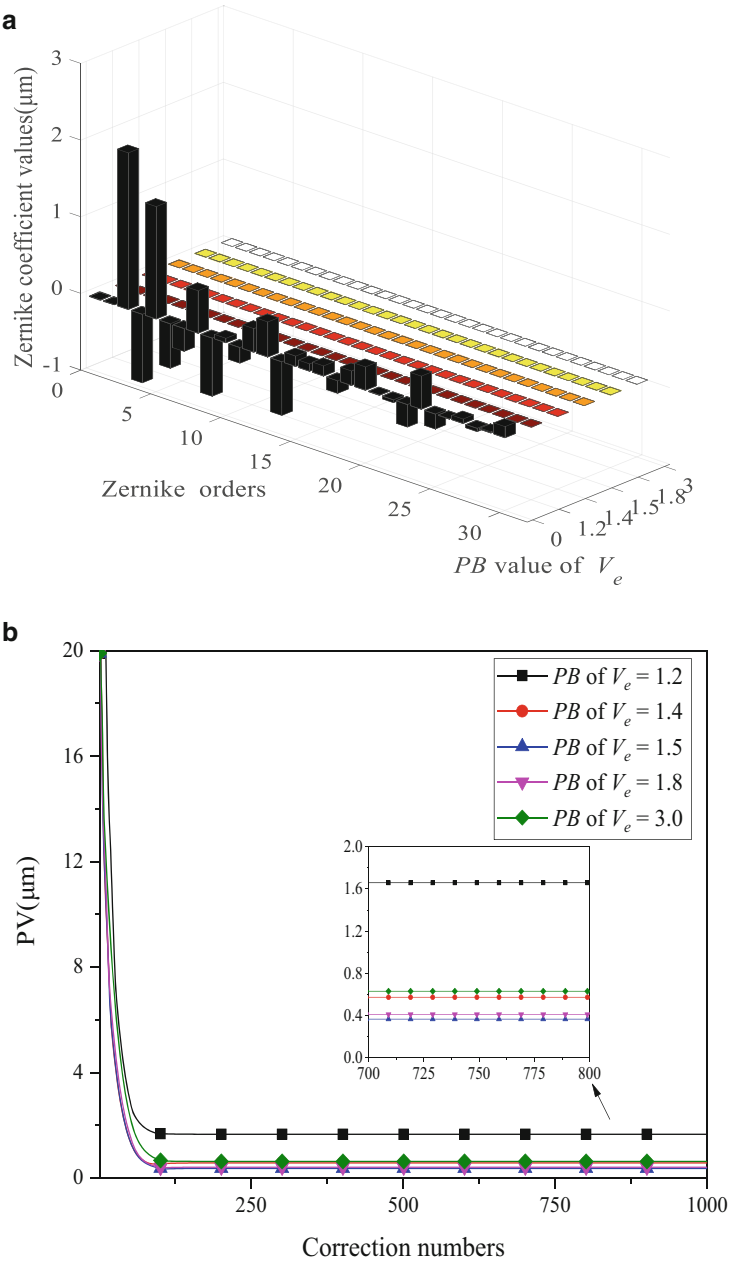


Fig. 24.6 Distribution of the Zernike coefficient value and PV value after correction of different control voltages [92] (**a**) Zernike coefficient values; (**b**) PV values

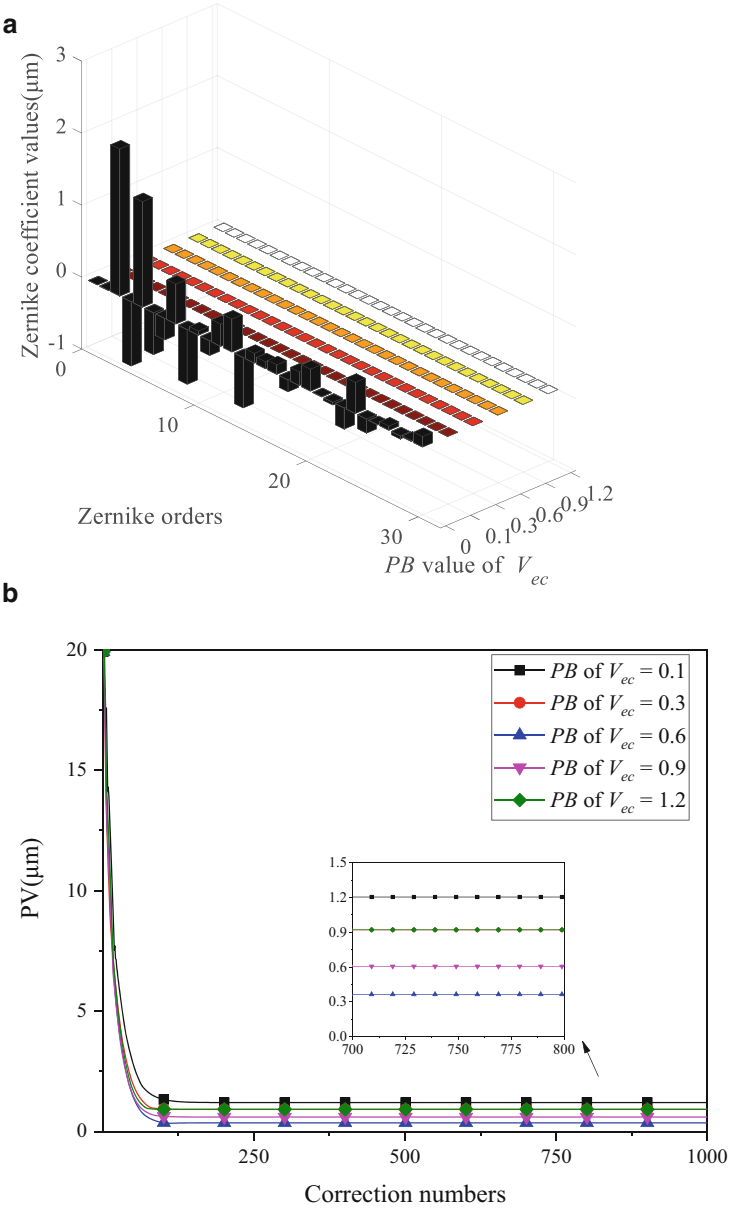


Fig. 24.7 Distribution of the Zernike coefficient value and PV value after correction of the first derivative of different control voltages [92] (**a**) Zernike coefficient values; (**b**) PV values

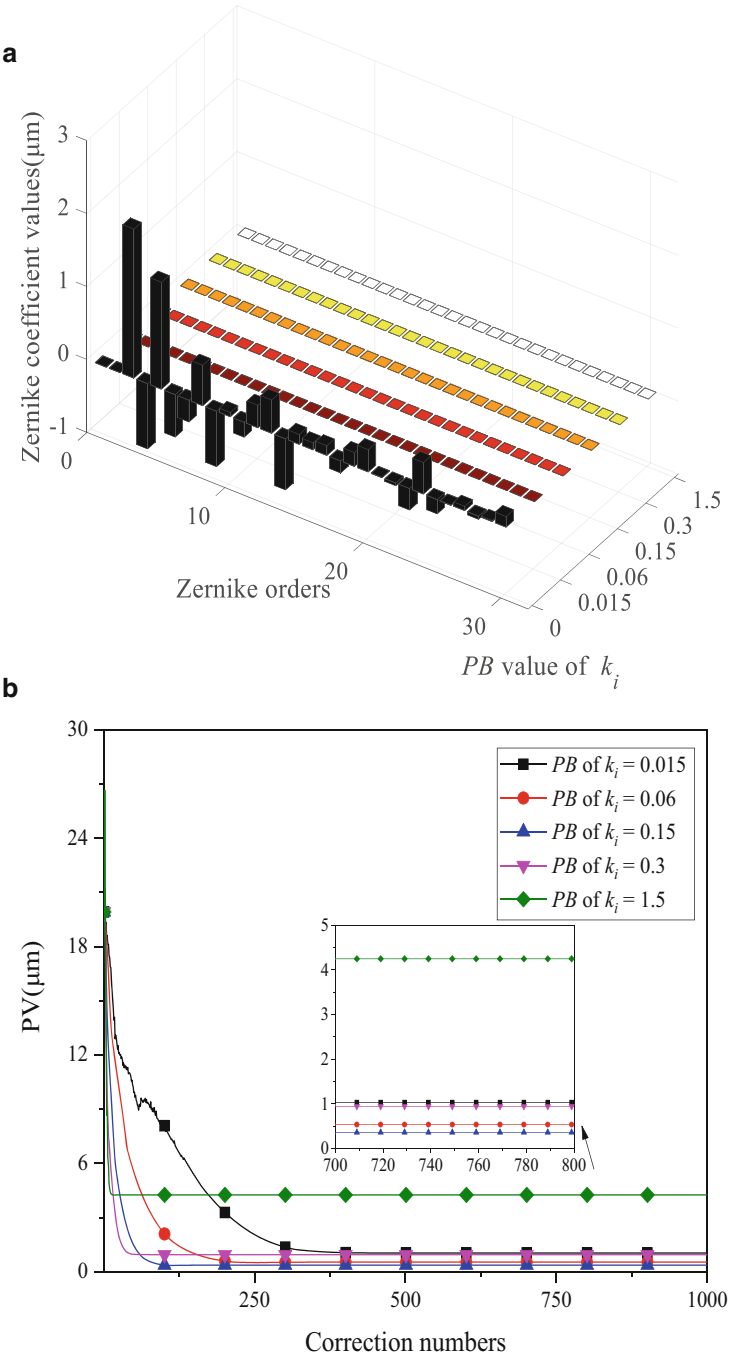


Fig. 24.8 Distribution of Zernike coefficient values and PV values after correction by different integration coefficients [92] (**a**) Zernike coefficient values; (**b**) PV values

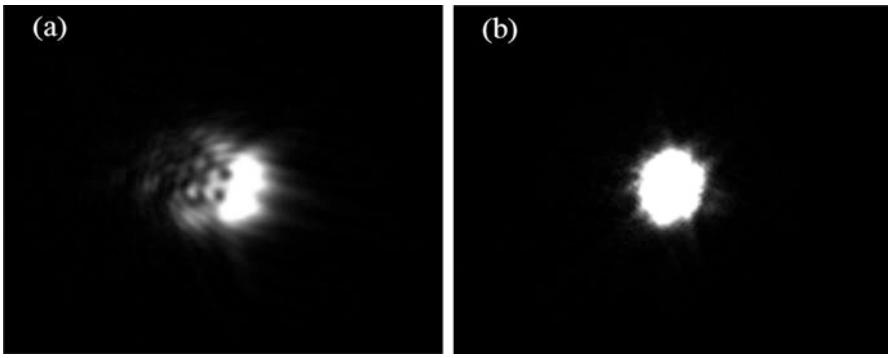


Fig. 24.9 Field-experiment light-spot diagram [62] (**a**) is the aberrated light spot, (**b**) is the light spot after improvement by WSAO system

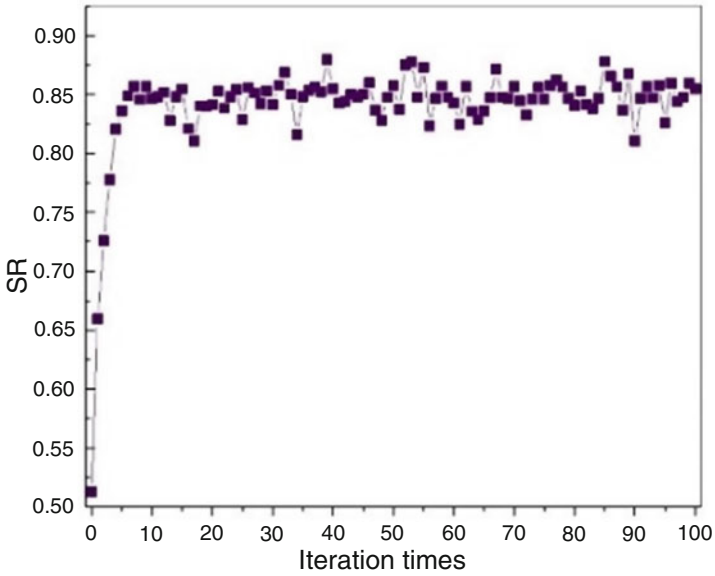


Fig. 24.10 SR change curve of field experiment [62]

1. When performing algorithm operations, software programs may require a long time. Hardware, such as digital signal processing (DSP) chips and ARM microprocessors, can be used to reduce the system running time and improve the system performance.
2. A model-free algorithm and a pattern method can be combined, and a new hybrid algorithm can be proposed for specific research.
3. The mode method includes a variety of base modes. Different modes can be applied to a system for research, to analyze the differences in the correction results of different modes, and improve these differences as much as possible.

In recent years, some researchers have used deep-learning techniques (such as residual networks, convolutional neural networks, and attention mechanisms) to obtain Zernike mode coefficients that conform to actual wavefront aberrations, and the obtained mode coefficients have high accuracy. Applying deep-learning technology to adaptive optics is a new research direction; however, the current related research is still insufficient and requires more in-depth study.

References

1. Rao Changhui, Jiang Wenhan. Analysis of the effectiveness of Adaptive Optical system for atmospheric turbulence compensation [J]. High Power Laser and Particle Beams, 1996, 8(4): 469–475.
2. Cubalchini R. Modal Wave-Front Estimation from Phase Derivative Measurements [J]. Opt. Soc. Am, 1979, 69(7): 972–977.
3. Babcock H W. The possibility of compensating atmospheric seeing [J]. Publications of the Astronomical Society of the Pacific, 1953, 65(386):229–236.
4. Noll R J. Zernike polynomials and atmospheric turbulence [J]. JOSA, 1976, 66(3): 207–211.
5. Fried D L. Adaptive optics topical issue [J]. Journal of the Optical Society of America A, 1977, 67(1): 426.
6. Jiang Wenhan, Zhang Yudong, Rao Changhui, et al. Research progress of adaptive optics in Institute of Optoelectronic Technology, Chinese Academy of Sciences [J]. Acta Optica Sinica, 2011, 31(09):72–80.
7. Nicolas A R. Atmospheric wavefront simulation using Zernike polynomials [J]. Optical Engineering, 1990, 29(10):1174–1181.
8. Daniel M H, Martin C V, Jose J S. Wavefront fitting with discrete orthogonal polynomials in a unit radius circle [J]. Optical Engineering, 1990, 29(06):672–676.
9. Hu Zhaohui, Jiang Wenhan. A preliminary study of eigenmode wavefront reconstruction [J]. Optical Engineering, 1995, 22(2):20–37.
10. Glenn D B, Christopher D. Zernike expansions for non-Kolmogorov turbulence [J]. Journal of the Optical Society of America A – Optics Image Science and Vision, 1996, 13(3):517–526.
11. Butts R R. The ABLE ACE wavefront sensor [J]. Physica A: Statistical Mechanics and its Applications, 1997, 3065(1):339–354.
12. Marcial M H, Manuel S, Daniel M H. Wavefront fitting using Gaussian functions [J]. Optics Communications, 1999, 163(4):259–269.
13. Neil M A, Booth M J, Wilson T. Closed-loop aberration correction by use of a modal Zernike wave-front sensor [J]. Optics Letters, 2000, 25(17):1083–1095.
14. Zhang Qiang, Lv Baida, Jiang Wenhan. Wavefront Reconstruction of Zernike Mode Method on Circular Regions [J]. High Power Laser and Particle Beams, 2000, 1(03): 306–310.
15. Hofer H, Artal P, Singer B, et al. Dynamics of the eye's wave aberration [J]. Journal of the Optical Society of America A – Optics Image Science and Vision, 2001, 18(3):497–506.
16. Chang Rong Seng, Sheu Jin Yi. Analysis of Seidel aberration by use of the discrete wavelet transform [J]. Applied Optics, 2002, 14(13):2408–2413.
17. Mahajan V N. Zernike polynomials and aberration balancing [J]. Physica A: Statistical Mechanics and its Applications, 2003, 5173(1):302–319.
18. Upton R, Ellerbroek B. Gram-Schmidt orthogonalization of the Zernike polynomials on apertures of arbitrary shape [J]. Optics Letters, 2004, 29(24):2840–2852.
19. Kawashima H, Nakanishi M, Takeda N, et al. Compensation of model eye's aberration by using deformable mirror [J]. Physica A: Statistical Mechanics and its Application, 2005, 5717(1): 219–229.
20. Seifert L, Tiziani H J, Osten W. Wavefront reconstruction with the adaptive Shack–Hartmann sensor [J]. Optics Communications, 2005, 245(1): 255–269.

21. Yang Huafeng, Jiang Zongfu. Reconstruction of 19-element Hartmann measurement wavefront by Zernike mode method [J]. Laser Technology, 2005, 29(5): 484–487.
22. Luo Zhifeng, Chen Huaixin, Ding Lei. Reconstruction of laser small scale distorted wavefront using Zernike method [J]. Laser Journal, 2006, 27(3): 35–36.
23. Booth M. Wave front sensor-less adaptive optics: a model-based approach using sphere packings [J]. Optics Express, 2006, 14(4): 1339–1352.
24. Fernandez E J, Vabre L, Hermann B, et al. Adaptive optics with a magnetic deformable mirror: applications in the human eye [J]. Optics Express, 2006, 14(20): 8900–8919.
25. Booth M J, Débarre D, Wilson T. Image-based wavefront sensorless adaptive optics [J]. Physica A: Statistical Mechanics and its Applications, 2007, 6711: 1102–1119.
26. Booth M J. Wavefront sensorless adaptive optics for large aberrations [J]. Optics Letters, 2007, 32(1): 5–7.
27. Rocha K M, Vabre L, Harms F, et al. Effects of Zernike wavefront aberrations on visual acuity measured using electromagnetic adaptive optics technology [J]. Journal of Refractive Surgery, 2007, 23(9): 953–959.
28. Durán V, Climent V, Tajahuerce E, et al. Efficient compensation of Zernike modes and eye aberration patterns using low-cost spatial light modulators [J]. Journal of Biomedical Optics, 2007, 12(1): 37–53.
29. Qi Bo, Chen Hongbin, Liu Shunfa. Regression Analysis Method of Zernike Polynomial Wavefront Fitting [J]. Optics and Precision Engineering, 2007, 17(3): 396–400.
30. Yang Ping, Xu Bing, Jiang Wenhan, et al. Application of genetic algorithm in adaptive optical system [J]. Acta Optica Sinica, 2007, 27 (9): 1628–1632.
31. Yang Ping, Ao Mingwu, Liu Yuan, et al. Adaptive Optical genetic algorithm based on Zernike mode coefficients [J]. Chinese Journal of Laser, 2008, 35(3): 367–372.
32. Yang Hui Zhen, Li Xinyang. Random parallel layer descent algorithm for adaptive optics system based on Zernike mode [J]. High Power Laser and Particle Beams, 2009, 21(5): 645–648.
33. Wang F. Wavefront sensing through measurements of binary aberration modes [J]. Applied optics, 2009, 48 (17): 2865-2870.
34. Liu Chao, Mu Quan Quan, Hu LiFa, et al. High porecision Zernike modal gray map reconstruction for liquid crystal corrector [J]. Chinese Physics B, 2010, 19(6): 14–26.
35. Mahajan V N, Aftab M. Systematic comparison of the use of annular and Zernike circle polynomials for annular wavefronts [J]. Applied optics, 2010, 49(33): 6489–6501.
36. Lee H. Use of Zernike polynomials for efficient estimation of orthonormal aberration coefficients over variable noncircular pupils [J]. Optics Letters, 2010, 35(13): 2173–2175.
37. Bueno J M, Acosta E, Schwarz C, et al. Wavefront measurements of phase plates combining a point-diffraction interferometer and a Hartmann-Shack sensor [J]. Applied Optics, 2010, 49(3): 450–456.
38. Chee O, Wang Zhaomin, Qu Weijuan, et al. A new method for improving the accuracy of wavefront fitting with Zernike polynomials [J]. Physics Procedia, 2011, 19(1): 134–138.
39. Rafael N, Ricardo R, Justiniano A. Representation of wavefronts in free-form transmission pupils with Complex Zernike Polynomials [J]. Journal of Optometry, 2011, 4(2): 41–48.
40. Yu Xuelian, Yao Yong, Sun Yunxu, et al. Stability of Zernike coefficients solved by the covariance matrix method in the analysis of the wave front aberration [J]. Optik – International Journal for Light and Electron Optics, 2010, 122(19): 1701–1706.
41. Feng Jie, Bai Yu, Xing Tingwen. Research on the fitting accuracy of Zernike polynomial wavefront [J]. Optical Instruments, 2011, 26(02): 31–34.
42. Wang Jianfeng, Liu Ke, Li Yanqiu. Comparison of wavefront reconstruction with modal method and zonal method for the inspection of catadioptric projection optics using Hartmann wavefront sensor [J]. Physica A: Statistical Mechanics and its Applications, 2011, 8197(1): 68–77.
43. Chen Bo, Chengchengqi, Pu Guoliang, et al. Wavefront Reconstruction of Generalized Ridge Estimation Mode Method for Image Restoration [J]. Infrared and Laser Engineering, 2012, 41(2): 478–483.

44. Booth M J, Débarre D, Jesacher A. Adaptive Optics for Biomedical Microscopy [J]. Optics and Photonics News, 2012, 23(1): 22–29.
45. Hasan S Y, Shaker A S. Study of Zernike polynomials of an elliptical aperture obscured with an elliptical obscuration [J]. Applied optics, 2012, 51(35):8490–8497.
46. I. Cacao, Morais J. An introduction to the quaternionic Zernike spherical polynomials [J]. AIP Conference Proceedings, 2013, 1758(1): 502–505.
47. Yang Huizhen, Wu Jian, Gong Chenglong. Model-based adaptive optics system without wavefront detection [J]. Acta Optica Sinica, 2014, 34(08): 17–21.
48. Yan Wei, Chen Zhihua, Du Taijiao, et al. Numerical Simulation of Atmospheric Thermal Halo Correction Based on Deformable Mirror Eigenmode Model Method [J]. Acta Optica Sinica, 2014, 34(11):9–18.
49. Yan Wei, Chen Zhihua, Du Taijiao, et al. Numerical Simulation of Adaptive Optics Correction for Thermal Halo Effect Mode Method [J]. Acta Optica Sinica, 2014, 34(09):9–19.
50. Yazdani R, Fallah H R, Hajimahmoodzadeh M. Reconstruction of two interfering wavefronts using Zernike polynomials and stochastic parallel gradient descent algorithm [J]. Optics Letters, 2014, 39(6):1705–1708.
51. Tian C, Chen X, Liu S. Modal wavefront reconstruction in radial shearing interferometry with general aperture shapes [J]. Optics express, 2016, 24(4): 3572–3583.
52. Wang Rui, Dong Bing. Adaptive optical system without wavefront sensor based on deformable mirror eigenmode for point target [J]. Chinese Journal of Lasers, 2016, 43(02): 221–233.
53. Zhang Xiang, Jiang Chuanfang, Hu Zhiqiu, et al. Mode characteristics analysis by using waverfront reconstruction for crystal Q-switched slab solid-state lasers [J]. Optik – International Journal for Light and Electron Optics, 2016, 127(6): 3316–3324.
54. Dong Bing, Li Yan, Han Xinli, et al. Dynamic Aberration Correction for Conformal Window of High-Speed Aircraft Using Optimized Model-Based Wavefront Sensorless Adaptive Optics [J]. Sensors, 2016, 16(9): 1414–1426.
55. Li Zhaokun, Zhao Xiaohui. BP artificial neural network based wave front correction for sensorless free space optics communication [J]. Optics Communications, 2017, 385: 219–233.
56. Wang Rui, Dong Bing. Square aperture laser beam cleaning based on deformable mirror eigenmode [J]. China Science paper, 2017, 12(05): 495–499.
57. Liu Ke, Wang Jiannian, Wang Hai, et al. Wavefront reconstruction for multi-lateral shearing interferometry using difference Zernike polynomials fitting [J]. Optics and Lasers in Engineering, 2018, 106(1): 75–81.
58. WANG Sheng-qian, WEI Kai, RAO Chang-hui. an adaptive optical device based on eigenmode control of pyramid wavefront sensor [P]. China:CN108646406A, 2018.
59. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84–90.
60. Ma Huimin, Liu Haiqiu, Qiao Yan, et al. Numerical study of adaptive optics compensation based on Convolutional Neural Networks [J]. Optics Communications, 2019, 433(1): 283–289.
61. Sun Wenhan, Wang Shuai, He Xing, et al. A Mode Restoration Method for Obtaining Uncoupled Zernike Coefficients in Transverse Shear Interferometry [J]. Opto-Electronic Engineering, 2019, 46(05): 13–27.
62. Ke Xizheng, Li Mei. Laser beam distorted wavefront correction based on deformable mirror eigenmodes [J]. Optical Engineering, 2019, 58(12): 6101–6119.
63. Ke Xizheng, Yang Ke, Zhang Ying. Experimental study on detecting wavefront distortion with CAPIS technology [J]. Laser & Optoelectronics Progress, 2019, 56(12): 30–37.
64. WuYang, Yang Huizhen, Xu Qi, et al. Convergence Speed Optimization of Model Adaptive Optics System Without Wavefront Detection [J]. Laser Journal, 2020, 41(03): 67–72.
65. Liang Jiaxin, Xiang Rujian, Du Yinglei, et al. Purification of beam of light based on deformable mirror eigenmode and far-field measurement [J]. High Power Laser and Particle Beams, 2020, 32(08): 10–19.

66. Rodin I A, Khonina S N, Serafimovich P G, et al. Recognition of wavefront aberrations types corresponding to single Zernike functions from the pattern of the point spread function in the focal plane using neural networks [J]. Computer Optics, 2020, 44(6): 923–930.
67. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278–2328.
68. Wu Yu, Guo Youming, Bao Hua, et al. Sub-millisecond phase retrieval for phase-diversity wavefront sensor [J]. Sensors, 2020, 20(1): 4877.
69. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [C]//Proceedings of the 18th International Conference. Munich, Germany, Springer, Cham, 2017: 234–241.
70. He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep Residual Learning for Image Recognition [J]. IEICE Transactions on Fundamentals of Electronics, Communications and C, 2016, 17(12): 770–778.
71. Hu Leijia, Hu Shuwen, Gong Wei, et al. Deep learning assisted Shack-Hartmann wavefront sensor for direct wavefront detection [J]. Optics Letters, 2020, 45(13): 3741–3744.
72. Wang Yukun, Jin Chengbin, Zhang Xingyun, et al. Wavefront Decoupling Method for Liquid Crystal Deformable Mirror Adaptive Optics System Based on Deformable Mirror Intrinsic Mode [P]. China: CN113311779A, 2024.
73. Cao Yang, Zhang Zupeng, Peng Xiaofeng, et al. Wavefront distortion correction method based on residual attention network [P]. China:CN114693938A, 2026.
74. Ning Yu, Zhang Yang, He Yu-long, et al. Adaptive optics wavefront correction method based on machine learning [P]. China:CN114488518A, 2026.
75. Wu Jiangli. Experimental research on coherent optical communication system without wavefront detection [D]. Xi'an: Xi'an University of Technology, 2018.
76. Cao J, Liu W, Zhao X, et al. Model-Based Tabu Search Algorithm for Free-Space Optical Communication with a Novel Parallel Wavefront Correction System [J]. Journal of the Optical Society of Korea, 2017, 19(1): 45–54.
77. Yu Ji, Dong Bing. Experimental research on wavefront sensorless adaptive optical system based on deformable mirror eigenmode [J]. Acta Optica Sinica, 2017, 35(3): 286–293.
78. Ji Zhongye. Research on high speed and high precision wavefront reconstruction in adaptive optics [D]. Beijing: Beijing University of Technology, 2018.
79. Wang Jiemei. Study on wavefront aberration based on Zernike polynomial [J]. Modern Computer, 2014, 1(2):46–49.
80. Yang Ke. Research on Computational Light Field Imaging Wavefront Sensing Technology [D]. Xi'an: Xi'an University of Technology, 2021.
81. Ke Xizheng, Zhang Yunfeng. Wavefront free sensing adaptive system and the method to improve the convergence speed by using the system: China. CN110365404A[P]. 2021.
82. Débarre D, Booth M J, Wilson T. Image based adaptive optics through optimization of low spatial frequencies [J]. Optical Express, 2007, 17(13): 8176–8190.
83. Lukosz W. Der Einfluß Der Aberrationen Auf Die Optischeübertragungsfunktion Bei KleinenOrts-Frequenze [J]. Optica Acta, 1963, 10(1):1–21.
84. Braat J. Polynomial Expansion Of Severely Aberrated Wave Fronts [J]. Opt. Soc. Am. A, 1987, 4(11):643–650.
85. Li Mei. Experimental study on eigenmode correction of beam wavefront distortion [D]. Xi'an: Xi'an University of Technology, 2023.
86. Ke Xizheng, Yang Shangjun, Wu Jia-li, et al. Research progress of adaptive optics in Optical Wireless Communication system of Xi'an University of Technology [J]. High Power Laser and Particle Beams, 2021, 33(08): 30–52.
87. Yu Ji, Dong Bing. Research on wavefront error correction of space optical remote sensor based on deformable mirror eigenmode [J]. Acta Optica Sinica, 2014, 34(12):308–314.
88. Singh J, Kumar N. Performance analysis of different modulation format on free space optical communication system [J]. Optik, 2013, 124(20): 4651–4654.
89. Liang Kunmiao. Mathematical physics method [M]. Beijing: People's Education Press, 1961.

90. Gao B, Li Q, He Y, et al. Absolute measurement of square flats with Legendre polynomial fitting method [C]//International Symposium on Optoelectronic Technology and Application 2014: Laser and Optical Measurement Technology and Fiber Optic Sensors. Beijing China, SPIE, 2014, 9297: 90–96.
91. Wang Kaidi. Study on adaptive optical correction of atmospheric turbulent wavefront distortion [D]. Xi'an: University of Chinese Academy of Sciences (Xi'an Institute of Optics and Precision Machinery, Chinese Academy of Sciences), 2014.
92. Ke Xizheng, Zhang Danyu. Fuzzy control algorithm for adaptive optical system [J]. Applied Optics, 2019, 58(36): 9967–9975.

Research on a Decoupling Algorithm for a Dual-Deformable-Mirror Wavefront-Correction System

25

Contents

Introduction 918
Domestic and Foreign Research Progress 919
Wavefront-Correction Principle with Dual-Deformable Mirrors 923
Principle of a Dual-Deformable-Mirror Correction System 923
Principle of the Separation Coefficient Algorithm 923
Numerical Simulation Analysis 930
Experimental Research 932
Coupling Analysis 935
Conclusion 937
References 938

Abstract

Wavefront distortion caused by atmospheric turbulence can be described as different types of aberrations, such as piston, tilt, defocusing, astigmatism, coma, etc. Dual-deformable mirrors can have mutual coupling effects that help correct the effects of wavefront distortion. In this study, a fast-steering mirror (FSM) and a deformable mirror (DM) were used to form a dual-deformable-mirror wavefront-correction system, and a decoupling algorithm that could correct any specified aberration was proposed. In this decoupling algorithm, both the FSM and the DM are controlled using the mode method, and specific aberrations are corrected based on a limited matrix. The DM-compensation ability is directly characterized by the mode coefficients of the aberrations; thus, it can independently correct any order of aberration and effectively reduce the coupling effect of the wavefront-correction system. An experiment using an adaptive-optical dual-deformable-mirror wavefront-correction system was conducted to verify the proposed decoupling algorithm. When the FSM corrects the first- and second-order aberrations, and the DM corrects the third-, tenth-, and 25th-order aberrations, the coupling coefficients are approximately 1.117×10^{-3}, 1.824×10^{-2}, and 7.81×10^{-3}, respectively, and the magnitude of the coupling coefficients

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_25

reaches 10^{-2} and below. The experimental results show that the decoupling algorithm can effectively suppress the coupling effect between the FSM and DM.

Keywords

Fast-steering mirror · Deformable mirror · Decoupling algorithm · Coupling coefficient

Introduction

The woofer–tweeter system is a typical dual-deformable-mirror wavefront-correction system [1] consisting of a large-stroke, low-spatial-frequency deformable mirror (woofer) and a small-stroke, high-spatial-frequency deformable mirror (tweeter). These two deformable mirrors work together to improve wavefront correction.

However, when both deformable mirrors work simultaneously, coupling effects occur, which can waste the correction ability of some deformable mirrors. If this phenomenon cannot be effectively suppressed, coupling errors will gradually accumulate; eventually, the coupling will completely consume the stroke of the deformable mirror. When the aberration to be corrected in the system changes, the deformable mirrors affected by coupling have difficulty effectively compensating for the aberration, leading to a decrease in the system correction ability [2]. Therefore, coupling compensation for deformable mirrors has become an important research direction for dual-deformable-mirror wavefront-correction systems, in terms of aberration-compensation ability and stability.

To fully utilize the advantages of adaptive-optics dual-deformable-mirror wavefront-correction systems, it is crucial to further improve their performance and eliminate the potential coupling between the deformable mirrors. Numerous scholars have addressed this issue, using different approaches proposed over the years. These decoupling algorithms have been successfully applied to various adaptive-optical dual-deformable-mirror wavefront-correction systems, significantly reducing the coupling effects between the mirrors.

Nonetheless, few decoupling algorithms are available for dual-deformable-mirror wavefront-correction systems that combine a fast-steering mirror (FSM) and a deformable mirror (DM), in which each aberration order cannot be independently controlled or corrected. Different application scenarios have varying requirements for different types of aberrations; hence, specific aberrations can be corrected according to the actual requirements to meet the diverse needs of different application scenarios, improve the system performance, and lay the foundation for large-aperture telescope applications.

This article proposes a decoupling algorithm to address the issues associated with the coupling between the FSM and DM in a dual-deformable-mirror wavefront-correction system. The algorithm utilizes the mode method to control both deformable mirrors, with the initial control voltage for the deformable mirror obtained

directly from the Zernike coefficient. The final control voltage of the deformable mirror is then determined using a proportional-integral (PI) controller, enabling the independent correction of any order aberration while avoiding cross-coupling effects. This approach significantly improves the performance and flexibility of the system, and can greatly enhance its capabilities.

Domestic and Foreign Research Progress

In 1990, Jiang et al. proposed a new algorithm based on a direct-slope gradient, derived a control decoupling matrix, and decoupled a deformable-mirror driver by adding appropriate fitting terms. The correctness of the algorithm was verified through computer simulations and experiments [3].

In 2002, Vorontsov et al. proposed a decoupled stochastic parallel gradient-descent optimization algorithm for adaptive optics, which achieved a fast convergence speed by partially decoupling the control channels of the adaptive system. The results showed that the phase distortion could be compensated with only 10–20 iterations [4].

In 2005, Hu et al. proposed a limited correction algorithm that decoupled the aberration of dual-deformable mirrors by solving the extended slope-response matrix of the deformable mirror. The numerical simulation of this algorithm showed that the correction effect of the dual-deformable-mirror adaptive-optical system was better than that of the finite-stroke single-deformable-mirror adaptive-optical system, and it avoided the production of deformable mirrors with both large-stroke and high-spatial-frequency characteristics [5]. Owing to the different response functions of each actuator, in 2009, Hu also proposed a decoupling algorithm based on an actual dual-deformable-mirror adaptive-optical system, verifying the feasibility of decoupling the two deformable mirrors in the system [6].

In 2006, Conan et al. proposed a control method based on distributed-mode decoupling to correct turbulence aberrations that increased with the increase of the telescope aperture. This method analyzed the mathematical characteristics of the influence function of wavefront correctors and solved for a series of distributed modes with good orthogonality between different wavefront correctors. It controlled the wavefront correctors to operate according to these specific modes, thereby suppressing the coupling between them and effectively eliminating coupling errors between different wavefront correctors [7].

In 2008, Zou et al. proposed a regional wavefront-control algorithm that used Lagrange-damping least-squares to suppress the correlation between two deformable mirrors [8].

In 2009, Ye et al. [9] proposed a closed-loop control method for a single-wavefront sensor and multiple deformable mirrors in an adaptive-optical system. It used an aberration-decoupling control algorithm to decompose the wavefront distortion and correct different deformable mirrors. The low-order deformable mirror (19-unit deformable mirror) used the mode method to correct low-order aberrations,

whereas the high-order deformable mirror (127-unit deformable mirror) used the direct-slope method to correct the residual aberrations.

A numerical simulation of the joint closed-loop operation of multiple deformable mirrors was conducted. A tilted mirror was used to correct wavefront tilt aberrations; 19 deformable mirrors were used to correct defocus aberrations; 61 deformable mirrors were used to correct Zernike aberrations before the 20th order, excluding tilt and defocus aberrations; and 127 deformable mirrors were used to correct other higher-order aberrations. The simulation results showed that this method enabled the joint closed-loop operation of multiple deformable mirrors, significantly improving the adaptive-optical correction ability. Multiple experiments were conducted using a dual-deformable-mirror adaptive-optical system. After multiple repeated experiments, the closed-loop working state of the system was stable and the repeatability was good, proving the effectiveness of the method and obtaining good applications in practical systems.

In 2010, Sredar et al. proposed a dual-deformable-mirror decoupling control method based on response-matrix orthogonalization. This method explored the coupling relationship between different deformable-mirror response matrices through mathematical derivation and used a matrix operation to remove the coupling components between the deformable-mirror response matrices, thereby orthogonalizing different deformable-mirror response matrices. This method was applied to a dual-deformable-mirror laser-scanning ophthalmic imaging system to verify its effectiveness [10].

In 2012, Lei et al. proposed a decoupling control algorithm based on mode decomposition. This method used the aberration-mode coefficient as the variable to be optimized and used different-order aberration-mode coefficients to control different deformable mirrors for aberration compensation, thereby decoupling the control of dual-deformable mirrors. The beam quality was better than that obtained using a single-deformable mirror for correction [11].

In 2013, Liu et al. proposed a decoupling control algorithm based on Zernike mode decomposition. The algorithm first decoupled the slope measured by the wavefront sensor using the Zernike mode, and then generated a control vector for the low-order deformable mirror using low-order Zernike coefficients to eliminate the high-order mode. In addition, the control vector of the high-order deformable mirror was reset through the constraint matrix to avoid accumulating coupling errors. Experiments showed that this algorithm effectively compensated for phase distortion and significantly suppressed the coupling between low- and high-order deformable mirrors [12].

In 2014, Li et al. proposed a constraint correction algorithm (CCA) for a combined deformable-mirror adaptive-optics (CDM-AO) system and analyzed its decoupling principle. The calibration results of this system were compared with those of traditional adaptive-optical systems. The simulation results showed that this system corrected different aberrations with different characteristics, had better calibration performance, and eliminated the coupling components of low-order deformable mirrors in high-order deformable mirrors. The coupling efficiency of this system was also analyzed and numerically evaluated, indicating that the algorithm

could improve the coupling efficiency of free-space optical (FSO) communication systems [13].

In 2015, Zhou et al. proposed an algorithm based on dual-deformable-mirror voltage decoupling. This algorithm used inertial components as controllers for low-order deformable mirrors and integral components as controllers for high-order deformable mirrors. Simulations and experimental results showed that this algorithm effectively corrected large-scale low-order aberrations and high-spatial-frequency aberrations in adaptive-optical systems. It significantly improved the correction efficiency, effectively reduced system residuals, and improved the performance of adaptive-optical systems [14].

In 2016, Liu et al. proposed a region-decoupling algorithm for controlling dual-deformable mirrors. The low-order deformable mirror is controlled using traditional slope-based correction methods, whereas the control vector of the high-order deformable mirror is generated by the response matrix of the residual wavefront and then reset through a constraint matrix to avoid coupling with the low-order deformable mirror. By comparing this algorithm with traditional regional-decoupling algorithms through simulations, the results showed that the proposed algorithm effectively suppressed coupling errors. The experimental system verified that this algorithm not only effectively compensated for phase distortion but also suppressed coupling errors and improved the system performance [2].

In 2018, Cheng et al. proposed a slope-based decoupling algorithm to simultaneously control two deformable mirrors in adaptive-optical systems. The basic principle of this algorithm was to directly use a Hartmann wavefront sensor to measure the response matrix of the low-order deformable mirror to construct a slope orthogonal basis. It then selectively allocated significant low-order aberrations to the low-order deformable mirror through the slope orthogonal basis and allocated the remaining aberrations to the high-order deformable mirror. By constructing a constraint matrix based on the response matrix and the slope orthogonal basis of the high-order deformable mirror, the control vector of the high-order deformable mirror was reset. Numerical simulation results showed that this algorithm fully utilized the compensation ability of low-order deformable mirrors, released the stroke of high-order deformable mirrors to compensate for high-order aberrations, and better suppressed coupling errors [15].

In the same year, to verify the effectiveness and practicality of the aforementioned algorithm, Cheng et al. established a typical dual-deformable-mirror adaptive-optical system. Theoretical model simulations were conducted for this system, and the results showed that the algorithm could selectively correct distorted wavefronts of different spatial frequencies for two deformable mirrors, effectively suppressing the cross-coupling between them [16].

In 2018, Cheng et al. proposed a projection-based decoupling algorithm. This algorithm utilized the response matrix of low-order deformable mirrors to construct a slope orthogonal basis and assigned different spatial-frequency aberrations to two deformable mirrors. Moreover, the control-voltage vector of the high-order deformable mirror was projected onto the slope orthogonal basis, and the coupling between the high-order and low-order deformable mirrors was calculated. The high-order

deformable mirror obtained the control-voltage vector after eliminating the coupling effect, thereby suppressing the cross-coupling between the two deformable mirrors. Correction experiments were conducted on static and dynamic aberrations. The results showed that the algorithm could effectively suppress the cross-coupling between the two deformable mirrors [17].

In 2020, Cheng et al. proposed a decoupling control algorithm to apply dual-deformable-mirror adaptive-optical systems to a generalized irregular-pupil region. This algorithm constructed a slope orthogonal basis based on the response matrix measured by the Hartmann sensor, and assigned aberrations of different spatial frequencies to the deformable mirror. It also used the slope orthogonal basis and improved response matrix of the higher-order deformable mirror to construct a constraint matrix to reset the control-voltage vector of the higher-order deformable mirror to remove the coupling effect between the higher-order and lower-order deformable mirrors. Numerical simulations were conducted to validate the algorithm, and showed that it better suppressed the cross-coupling between low-order and high-order deformable mirrors [18].

To solve the collaborative control problem of dual wavefront correctors in generalized irregular-pupil regions, Kong et al. proposed a decoupling control algorithm based on numerical orthogonal polynomials (NOP) in 2021. Through numerical simulations and experiments, it was shown that this algorithm eliminated cross-coupling between dual-wavefront correctors and exhibited better performance [19].

In 2021, Yang et al. used a dual-deformable-mirror combination control for fiber power coupling to improve the efficiency of space optical-fiber coupling [20]. In the same year, Wang et al. proposed a global-decoupling control algorithm to achieve high-precision control of two correctors simultaneously. This algorithm utilized the eigenmode orthogonal basis and response matrix of the deformable mirror to construct a constraint matrix for the two correctors, thereby constructing a global-decoupling control matrix. Numerical simulations showed that this algorithm performed well in simultaneously correcting different spatial-frequency aberrations and suppressing cross-coupling between dual correctors, fully utilizing and improving the correction ability of deformable mirrors [21].

In 2022, Wang et al. proposed an intrinsic-mode wavefront-decoupling algorithm to control the decoupling of dual correctors in liquid-crystal deformable-mirror adaptive-optical systems and prevent the loss of correction ability caused by coupled aberrations. The simulation and experimental results showed that, after correcting 55 intrinsic modes, the root-mean-square of the residual aberration was 0.09 μm. After decoupling correction, the root-mean-square value of the residual aberration decreased by 0.02 μm. At 0.7–0.95 μm, the uncoupled imaging resolution in the visible band was 11.04 μm. After decoupling, the imaging resolution was 9.84 μm, which was approximately 1.1 and one times the diffraction-limit resolution, respectively [22].

In 2023, Liu et al. proposed a dual-deformable-mirror decoupling control method based on pattern-projection suppression, which eliminated the low-order correction components in high-order deformable mirrors through projection suppression, thereby enabling efficient collaboration between the two deformable mirrors. Theoretical analysis and simulation results showed that, compared with traditional

decoupling algorithms for wavefront-sensorless adaptive-optical systems, this decoupling control method had good coupling suppression and aberration correction and high practical engineering value [23].

A summary of the above research progress was made, and the research progress of dual-deformable-mirror decoupling technology is presented in Table 25.1.

Wavefront-Correction Principle with Dual-Deformable Mirrors

Principle of a Dual-Deformable-Mirror Correction System

The working principle of the dual-deformable-mirror correction system is illustrated in Fig. 25.1. The system consists of three parts: a wavefront sensor, wavefront controller, and two wavefront correctors [24]. When a light beam propagates in the atmosphere, the wavefront is distorted by atmospheric turbulence. It is corrected using a wavefront corrector.

In the dual-deformable-mirror correction system used in this study, the wavefront distortion was graded and corrected by two wavefront correctors. Wavefront corrector 1 is generally an FSM or a large-stroke, low-spatial-frequency DM (commonly referred to as a woofer) responsible for correcting low-order aberrations. Wavefront corrector 2 is generally a small-stroke, high-spatial-frequency DM (commonly referred to as a tweeter), which is responsible for correcting high-order aberrations [1].

The system utilizes wavefront sensors to measure the wavefront of the received beam in real time and feeds the measured wavefront information back to the wavefront controller. The controller calculates the control voltage of the wavefront corrector using a control algorithm and provides real-time control signals for the wavefront corrector. The wavefront corrector receives a real-time control voltage from the wavefront controller, fits the surface shape to correct the wavefront distortion, and improves the communication quality of the dual-deformable-mirror wavefront-correction system.

Principle of the Separation Coefficient Algorithm

The distorted wavefront caused by the interference of atmospheric turbulence can be expanded using Zernike polynomials. The expression for wavefront $\phi(\rho)$ is [25, 26]

$$\phi(\rho) = \sum_{i=1}^{\infty} a_i Z_i(\rho), \tag{25.1}$$

where a_i represents the ith-order Zernike coefficient and $Z_i(\rho)$ represents the ith-order Zernike polynomial [26]. This study uses the Zernike polynomial defined by Noll in polar coordinates, whose expression is [27]

Table 25.1 Research progress in dual-deformable-mirror decoupling technology

Year	Scholars	Research progress
1990	Jiang Wenhan et al.	A new algorithm based on a direct-slope gradient was proposed, and a control decoupling matrix was derived to verify the effectiveness of the algorithm
2002	Vorontsov et al.	A decoupled random parallel gradient-descent optimization algorithm for adaptive optics was proposed, which quickly converged
2005	Hu Shijie et al.	A limited correction algorithm was proposed to decouple the aberrations of dual-deformable mirrors by solving the extended slope-response matrix of the deformable mirror
2006	Conan et al.	A control method based on distributed-mode decoupling was proposed, which effectively eliminated coupling errors between different wavefront correctors
2008	Zou Weiyao et al.	A regional wavefront-control algorithm was proposed that utilized the Lagrange-damping least-squares method to suppress the correlation between two deformable mirrors
2009	Hu Shijie et al.	A decoupling algorithm based on an actual dual-deformable-mirror adaptive-optical system was proposed, and the feasibility of decoupling two deformable mirrors in the system was verified
2009	Ye Weihong et al.	A closed-loop control method with a single-wavefront sensor and multiple deformable mirrors for adaptive-optical systems was been proposed; its effectiveness was demonstrated and applied to practical systems
2010	Sredar et al.	A dual-deformable-mirror decoupling control method based on response-matrix orthogonalization was proposed and applied to a dual-deformable-mirror laser-scanning ophthalmic imaging system, and the effectiveness of the method was verified
2012	Lei Xiang et al.	A decoupling control algorithm based on pattern decomposition was proposed to control the decoupling of dual-deformable mirrors; the beam quality was better than using only a single-deformable mirror for correction
2013	Liu Wenjin et al.	A decoupling control algorithm based on Zernike mode decomposition was proposed, which effectively compensated for phase distortion and significantly suppressed the coupling between low-order and high-order deformable mirrors
2014	Li Zhaokun et al.	A constraint-correction algorithm was proposed, which eliminated the coupling components of low-order deformable mirrors in high-order deformable mirrors
2015	Zhou Rui et al.	A voltage-decoupling algorithm based on dual-deformable mirrors was proposed, which significantly improved the correction efficiency, effectively reduced the system residual, and improved the performance of the adaptive-optical system
2016	Liu Wenjin et al.	A region-decoupling algorithm was proposed to control dual-deformable mirrors. This algorithm not only effectively compensated for phase distortion, but also suppressed coupling errors and improved the system performance
2018	Cheng Tao et al.	A slope-based decoupling algorithm was been proposed, which had better performance in suppressing coupling errors

(continued)

Table 25.1 (continued)

Year	Scholars	Research progress
2018	Cheng Tao et al.	A projection-based decoupling algorithm was proposed, which effectively suppressed cross-coupling between two deformable mirrors
2020	Cheng Tao et al.	A decoupling control algorithm was proposed to apply dual-deformable-mirror adaptive-optical systems in the generalized irregular-pupil region
2021	Kong Lingxi et al.	A decoupling control algorithm based on numerical orthogonal polynomials was proposed, which eliminated cross-coupling between dual wavefront correctors and had better performance
2021	Wang Yukun et al.	A global-decoupling control algorithm was been proposed to simultaneously achieve high-precision control of two correctors
2022	Wang Yukun et al.	An eigenmode wavefront-decoupling algorithm was proposed to control the decoupling of dual correctors in liquid-crystal deformable-mirror adaptive-optical systems, preventing the loss of correction ability caused by coupled aberrations
2023	Liu Wenjin et al.	A dual-deformable-mirror decoupling control method based on pattern-projection suppression was proposed. Compared with traditional decoupling algorithms for wavefront-sensorless adaptive-optical systems, this decoupling control method had good coupling suppression and aberration correction, and high practical engineering value

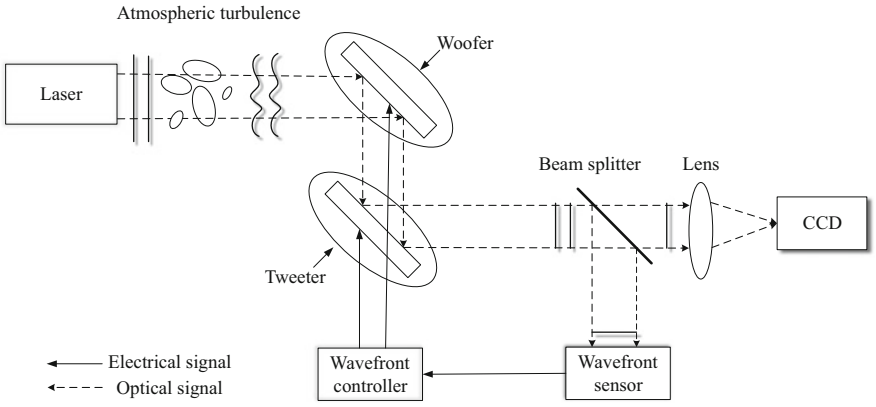


Fig. 25.1 Schematic diagram of a dual-deformable-mirror wavefront-correction system

$$Z_i = \begin{cases} \sqrt{n+1}R_n^m(\rho)\sqrt{2}\cos(m\theta) & m \neq 0 \text{ and } m \text{ is even} \\ \sqrt{n+1}R_n^m(\rho)\sqrt{2}\sin(m\theta) & m \neq 0 \text{ and } m \text{ is odd} \\ \sqrt{n+1}R_n^m(\rho) & m = 0 \end{cases}, \tag{25.2}$$

where (ρ, θ) is the polar coordinate. The integers m and n represent the angular and radial levels, respectively, and the relationship between them is as follows:

$$\begin{cases} m \le n \\ n - |m| = k \quad k \in N \end{cases}. \tag{25.3}$$

The radial polynomial $R_n^m(\rho)$ is defined as [28]

$$R_n^m(\rho) = \sum_{s=0}^{\frac{n-m}{2}} \frac{(-1)^s (n-s)!}{s![(n+m/2)-s]![(n-m/2)-s]!} \rho^{n-2s}. \tag{25.4}$$

The lower-order modes of the Zernike polynomials in the circular domain are consistent with the Seidel aberrations, and the specific expressions for the aberrations represented by the first 15 orders of the Zernike polynomials are listed in Table 25.2.

Assuming that the influence function of the ith actuator of the DM is f_i, it can generally be measured using a Zygo interferometer or wavefront sensor. The jth-order Zernike mode is denoted as Z_j. The cross-correlation matrix C_{fz} between the influence function of the DM and the Zernike modes can be represented as [23]

$$C_{fz}(i,j) = \int f_i(r) Z_j(r) d^2 r. \tag{25.5}$$

Similarly, the expression for calculating the autocorrelation matrix C_{ff} between the influence functions of the DM actuator is as follows [23]:

$$C_{ff}(i,j) = \int f_i(r) f_j(r) d^2 r, \tag{25.6}$$

where C_{ff} is the autocorrelation matrix of the influence functions of the deformable-mirror actuator.

According to Eqs. (25.5) and (25.6), the conversion matrix T from the Zernike mode coefficient to the control voltage of the DM can be obtained as [23]

$$T = C_{ff}(i,j)^{-1} \cdot C_{fz}(i,j), \tag{25.7}$$

where C_{fz} is the cross-correlation matrix obtained from Eq. (25.5). C_{ff} is the autocorrelation matrix obtained from Eq. (25.6). C_{ff}^{-1} is the inverse of matrix C_{ff}; then, the expression of the control voltage to be loaded each time by the DM is [23]

$$v = T \cdot a, \tag{25.8}$$

where $a = [a_1, a_2, \ldots, a_n]^T$, the dimension is $n \times 1$, and n is the order of the Zernike polynomial. T is the transformation matrix between Zernike mode coefficients and

Table 25.2 Aberration expressions for the first 15 orders of Zernike polynomials

i	(n, m)	Polar coordinates	Cartesian coordinates	Aberration type
1	(0, 0)	1	1	Piston
2	(1, 1)	$2\rho\cos\theta$	$2x$	X-direction tilt
3	(1, −1)	$2\rho\sin\theta$	$2y$	Y-direction tilt
4	(2, 0)	$\sqrt{3}(2\rho^2-1)$	$\sqrt{3}(2x^2+2y^2-1)$	Defocus
5	(2, −2)	$\sqrt{6}\rho^2\sin 2\theta$	$2\sqrt{6}xy$	Y-direction coma
6	(2, 2)	$\sqrt{6}\rho^2\cos 2\theta$	y^2-x^2	X-direction coma
7	(3, −1)	$2\sqrt{2}(3\rho^3-2\rho)\sin 2\theta$	$2\sqrt{2}(3x^3+3xy^2-2x)$	Y-direction astigmatism
8	(3, 1)	$2\sqrt{2}(3\rho^3-2\rho)\cos 2\theta$	$2\sqrt{2}(3x^3+3xy^2-2y)$	X-direction astigmatism
9	(3, −3)	$2\sqrt{2}\rho^3\sin 3\theta$	$2\sqrt{2}(3xy^2-x^3)$	Trefoil aberration
10	(3, 3)	$2\sqrt{2}\rho^3\cos 3\theta$	$2\sqrt{2}(y^3-3x^2y)$	Trefoil aberration
11	(4, 0)	$\sqrt{5}(6\rho^4-6\rho^2+1)$	$\sqrt{5}(6y^4+12x^2y^2+6x^4-6x^2-6y^2+1)$	Spherical aberration
12	(4, 2)	$\sqrt{10}(4\rho^4-3\rho^2)\cos 2\theta$	$\sqrt{10}(4y^4-4x^4+3x^2-3y^2)$	Spherical aberration
13	(4, −2)	$\sqrt{10}(4\rho^4-3\rho^2)\sin 2\theta$	$\sqrt{10}(8xy^3+8x^3y-6xy)$	Spherical aberration
14	(4, 4)	$\sqrt{10}\rho^4\cos 4\theta$	$\sqrt{10}(x^4-6x^2y^2+y^4)$	Quatrefoil aberration
15	(4, −4)	$\sqrt{10}\rho^4\sin 4\theta$	$\sqrt{10}(4xy^3-4x^3y)$	Quatrefoil aberration

DM control voltages. $v = [v_1, v_2, \ldots, v_P]^T$ is the initial control voltage of the DM and P is its number of actuators.

The Shack–Hartmann wavefront sensor (SH-WFS) can measure the slope; its slope matrix is represented as

$$S = Z \cdot a, \tag{25.9}$$

where $S = [S_{x1}, S_{x2}, \ldots, S_{xm}, S_{y1}, S_{y2}, \ldots, S_{ym}]$ is the slope matrix directly measured by the SH-WFS. The dimension is $2m \times 1$ and m is the number of subapertures of the SH-WFS. Z is the wavefront-reconstruction matrix and the dimension is $2m \times n$. $a = [a_1, a_2, \ldots, a_n]^T$, the dimension is $n \times 1$, and n represents the order of the Zernike polynomial.

According to the wavefront-reconstruction principle of the Zernike mode method [29], the Zernike coefficients can be reconstructed using the measured slope S, as follows:

$$a = Z^{+} \cdot S, \tag{25.10}$$

where Z^{+} is the generalized inverse matrix of Z.

The mode-coefficient expression for correcting the aberrations with a low-order DM is as follows:

$$a_w = I_w \cdot a = I_w \cdot Z^{+} \cdot S, \tag{25.11}$$

where $a_w = [a_{w1}, a_{w2}, \ldots, a_{wn}]^T$ is the mode coefficient assigned to the low-order DM for aberration correction. I_w is an $n \times n$ diagonal matrix, which is also known as the limiting matrix for low-order aberrations. For example, if the low-order DM corrects first- and second-order aberrations, the expression for I_w is as follows:

$$I_w = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 0 & & & \\ & & & \cdots & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix}. \tag{25.12}$$

Based on Eq. (25.8), the initial control voltage of the low-order DM can be directly calculated. The expression is

$$v_w = T_w \cdot a_w = T_w \cdot I_w \cdot Z^{+} \cdot S, \tag{25.13}$$

where $v_w = [v_{w1}, v_{w2}, \ldots, v_{wh}]^T$ is the final control voltage of the low-order DM, and h is its number of actuators. T_w is the conversion matrix between the Zernike mode coefficient and the control voltage of the low-order DM, as calculated by Eqs. (25.5), (25.6), and (25.7). a_w is the mode coefficient, calculated using Eq. (25.11), that the low-order DM will use to correct aberrations.

After obtaining the initial control voltage of the low-order DM using Eq. (25.13), a PI controller is used to solve the final control voltage of the low-order DM and send it to the low-order DM for operation. To accurately calculate the control voltage of the low-order DM, its final control-voltage expression at time k is

$$V_w(k) = k_{i_w} \cdot V_w(k-1) + k_{p_w} \cdot v_w(k-1), \tag{25.14}$$

where k_{i_w} is the integral-gain coefficient of the PI controller, and k_{p_w} is the proportional-gain coefficient of the PI controller. $V_w(k) = [V_{w1}, V_{w2}, V_{w3}, \ldots, V_{wh}]^T$ $V_{w2}, V_{w3}, \ldots, V_{wh}]^T$ is the final control voltage of the low-order DM and h is its

number of actuators. $v_w(k-1)$ is the previously calculated initial control voltage of the low-order DM.

In addition, the mode-coefficient expression for correcting aberrations in the high-order DM is as follows:

$$a_t = I_t \cdot a = I_t \cdot \mathrm{Z}^{+} \cdot S, \tag{25.15}$$

where $a_t = [a_{t1}, a_{t2}, \ldots, a_{tn}]^T$ is the mode coefficient assigned to the high-order DM for aberration correction. This is the $n \times n$ diagonal matrix, also called the limiting matrix of higher-order aberrations. For example, if a high-order DM only corrects third-order aberrations, the expression for I_t is as follows:

$$I_t = \begin{bmatrix} 0 & & & & & \\ & 0 & & & & \\ & & 1 & & & \\ & & & \cdots & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix}. \tag{25.16}$$

According to Eq. (25.8), the initial control voltage of the high-order DM can be directly calculated; its expression is

$$v_t = T_t \cdot a_t = T_t \cdot I_t \cdot \mathrm{Z}^{+} \cdot S, \tag{25.17}$$

where $v_t = [v_{t1}, v_{t2}, \ldots, v_{tt}]^T$ is the final control voltage of the high-order DM, and t is its number of actuators. T_t is the conversion matrix between the Zernike mode coefficient and the high-order DM control voltage, calculated using Eqs. (25.5), (25.6) and (25.7). a_t is the mode coefficient of the high-order DM calculated using Eq. (25.15), which corrects for aberrations.

After obtaining the initial control voltage of the high-order DM using Eq. (25.17), a PI controller is used to solve the final control voltage of the high-order DM and send it to the high-order DM for operation. To accurately calculate the control voltage of the high-order DM, its final control-voltage expression at time k is

$$V_t(k) = k_{i_t} \cdot V_t(k-1) + k_{p_t} \cdot v_t(k-1), \tag{25.18}$$

where k_{i_t} is the integral-gain coefficient of the PI controller, and k_{p_t} is the proportional-gain coefficient of the PI controller. $V_t(k) = [V_{t1}, V_{t2}, V_{t3}, \ldots, V_{tt}]^T$ is the final control voltage of the high-order DM and t is its number of actuators. $v_t(k-1)$ is the initial control voltage of the high-order DM calculated at the previous moment.

Using Eqs. (25.14) and (25.18), the final control voltages of the low-order and high-order DMs can be calculated to obtain the corrected wavefront phase. The corrected wavefront phase is denoted as ϕ^k, and the iterative phase expression after the kth correction is

$$\phi^{k+1} = \phi^k + \left(V_w^k \cdot f_w\right) + \left(V_t^k \cdot f_t\right), \tag{25.19}$$

where f_w is the influence function of the low-order DM, and f_t is the influence function of the high-order DM.

In practical applications, the measurement and calculation of Z, Z^+, T_w, T_t, f_w, and f_t can be completed in advance and saved for future use, allowing for quick and efficient access during the calibration process. This decoupling algorithm not only streamlines the program's performance by reducing its running time and storage space but also significantly enhances the computational efficiency. This decoupling algorithm provides the flexibility to independently correct any order of aberration, while also enabling the independent correction of each order of aberration. This makes it highly adaptable to different types of dual-deformable-mirror systems and enables efficient and accurate wavefront correction for a wide range of applications.

The peak-to-valley (PV) value represents the difference between the maximum and minimum values of the wavefront phase, indicating the maximum level of wavefront-phase defects. However, this approach ignores the proportion of this part of the overall wavefront. Therefore, evaluating the wavefront quality based solely on the PV value is not sufficiently rigorous. However, the root-mean-square (RMS) value focuses more on the overall wavefront-phase information. It reflects the degree of deviation of all the points on the wavefront from their ideal positions. Therefore, combining the PV and RMS values can more accurately represent the wavefront condition.

This study used PV and RMS values to evaluate the level of wavefront distortion [30]. Based on the phase compensation of the two deformable mirrors, the coupling coefficient (c) was used to quantitatively describe the coupling effect of the adaptive-optics dual-deformable-mirror correction system [12]. A c value closer to 0 indicates a smaller coupling effect between the woofer and tweeter, whereas a c value closer to 1 indicates a larger coupling effect.

Numerical Simulation Analysis

Experimental data were analyzed in the laboratory of Xi'an University of Technology to further validate the algorithm proposed in this study. The transmitting end of the data measurement was positioned in Building 5 of Xi'an University of Technology, whereas the receiving end was located in Building 2. The slope data of the wavefront distortion were measured using a Shack–Hartmann wavefront sensor.

According to the wavefront-reconstruction principle [29], the wavefront-reconstruction matrix was inverted and the product of the slope data measured by the Hartmann wavefront sensor was obtained (i.e., Eq. 25.10) to obtain the initial Zernike coefficient of the wavefront distortion, as shown in Fig. 25.2a. Given the Zernike-coefficient value of the wavefront distortion, the phase of the wavefront distortion can be calculated by substituting it into Eq. (25.1), as shown in Fig. 25.2b, which displays the phase diagram of the wavefront distortion.

Next, a dual-deformable-mirror wavefront-correction system consisting of an FSM (woofer) and a 69-unit DM (tweeter) was used to correct the wavefront-

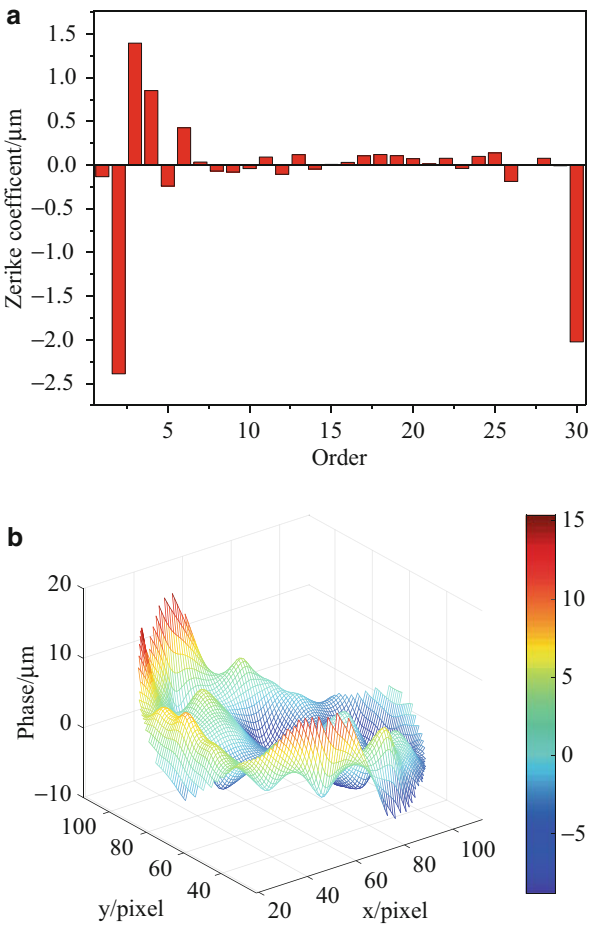


Fig. 25.2 Initial wavefront distortion. (**a**) Zernike coefficient; (**b**) wavefront-distortion phase diagram

distortion data measured in the experiment. The FSM corrected the second-order aberration, the DM corrected the thirtieth-order aberration, and the parameters of the PI controller were set to $k_{i_w} = 0.999$, $k_{p_w} = 0.1$, $k_{i_t} = 0.999$, and $k_{p_t} = 0.1$. Figure 25.3 shows the wavefront-phase diagrams corrected by the FSM and DM, as well as the wavefront-phase diagrams corrected by working together.

Figure 25.3 shows the phase diagrams of the FSM and DM after correcting the aberrations individually and in combination. Specifically, Fig. 25.3a illustrates the wavefront-phase diagram after correction with the FSM, whereas Fig. 25.3b shows the residual wavefront-phase diagram after correction with the DM. Figure 25.3c depicts the residual wavefront-phase diagram after both the FSM and DM were employed for correction.

As shown in Fig. 25.2a, the second- and thirtieth-order aberrations had the largest Zernike coefficients, indicating that they contributed significantly to the total

wavefront aberration. Therefore, the FSM was used to correct the second-order aberration, whereas the DM corrected the thirtieth-order aberration.

By comparing Fig. 25.2b with Fig. 25.3b, it is evident that the tilt degree of the corrected wavefront-distortion phase diagram is minimal, indicating that the residual wavefront after correction has a minimal proportion of tilt aberrations. However, when the DM was used to correct the thirtieth-order aberration without compensating for the second-order aberration, as shown in Fig. 25.3b, there was a degree of tilt in the corrected wavefront-phase diagram. As shown in Fig. 25.3c, there is a significant difference between the corrected wavefront-phase diagram and the initial distorted wavefront-phase diagram. By evaluating the PV and RMS values, it can be concluded that the wavefront-correction ability of the proposed algorithm is accurately reflected, demonstrating that the decoupling algorithm independently corrects the specified aberrations.

In a dual-deformable-mirror wavefront-correction system, the deformation generated by two deformable mirrors is typically used to correct the wavefront aberration. However, when the two deformable mirrors are coupled, their deformations can affect each other, leading to an unwanted coupling effect. To evaluate this effect, the coupling coefficient is employed as an important indicator.

The correction amounts of the FSM and DM were calculated based on the FSM and DM control voltage. By substituting these correction amounts into the coupling-coefficient expression [12], the value of the coupling coefficient between the FSM and DM can be solved.

The coupling-coefficient diagram is presented in Fig. 25.4, where the black line represents the coupling-coefficient curve for second-order aberrations corrected by the FSM and the second- and thirtieth-order aberrations corrected by the DM. It is observed that a coupling effect exists between the two mirrors with a coupling coefficient of about 4.8×10^{-1}, indicating an order of about 10^{-1}.

In the proposed algorithm, a finite matrix was used to constrain the Zernike coefficients. By modifying the elements on the diagonal of the matrix defined by Eqs. (25.12) and (25.16), it was ensured that only the second-order aberration was corrected by the FSM, whereas only the thirtieth-order aberration was corrected by the DM. As shown by the red line, this results in a coupling coefficient value around 10^{-5}, indicating that the algorithm effectively suppressed the coupling effect between the FSM and DM.

Experimental Research

In adaptive-optics systems, the FSM is usually regarded as a special DM that only compensates for tilt aberrations. To verify the importance of the theoretical results, an experimental platform for a dual-deformable-mirror wavefront-correction system was built indoors to correct the wavefront distortion. The experimental setup is illustrated in Fig. 25.5.

The beam was emitted by a 650-nm semiconductor laser and extended and collimated through a 4f system composed of lenses L_1 and L_2. The collimated

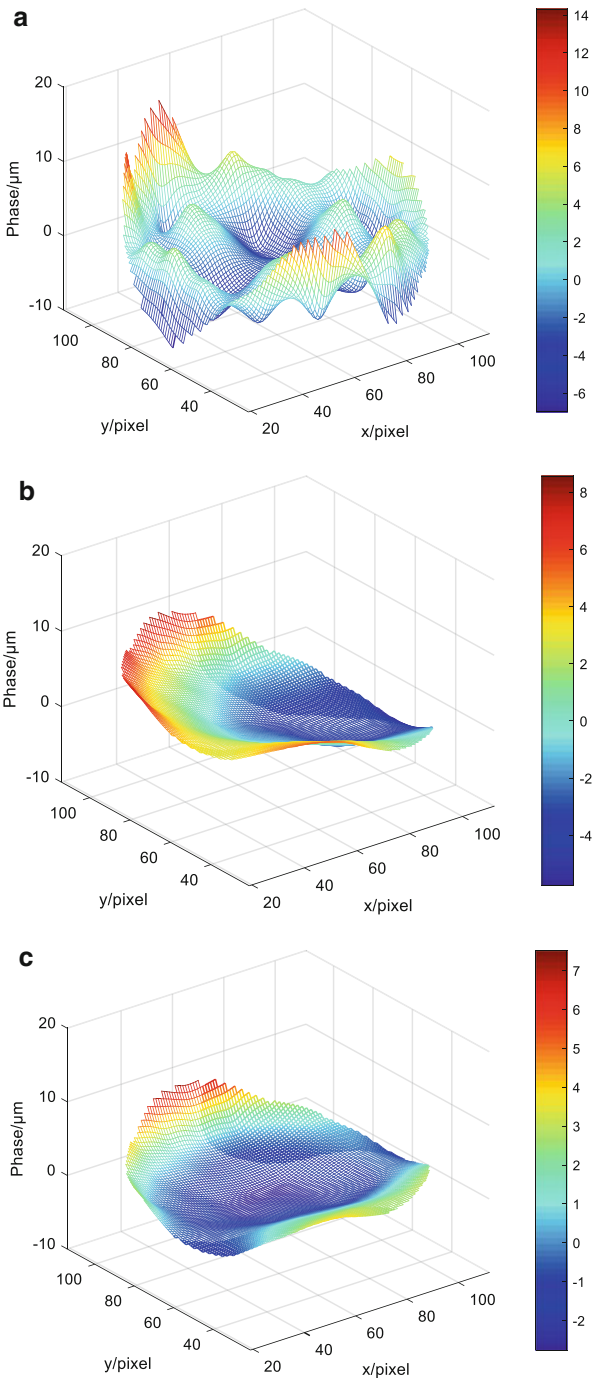


Fig. 25.3 Wavefront-phase diagrams. (**a**) After FSM correction; (**b**) after DM correction; (**c**) after both FSM and DM correction

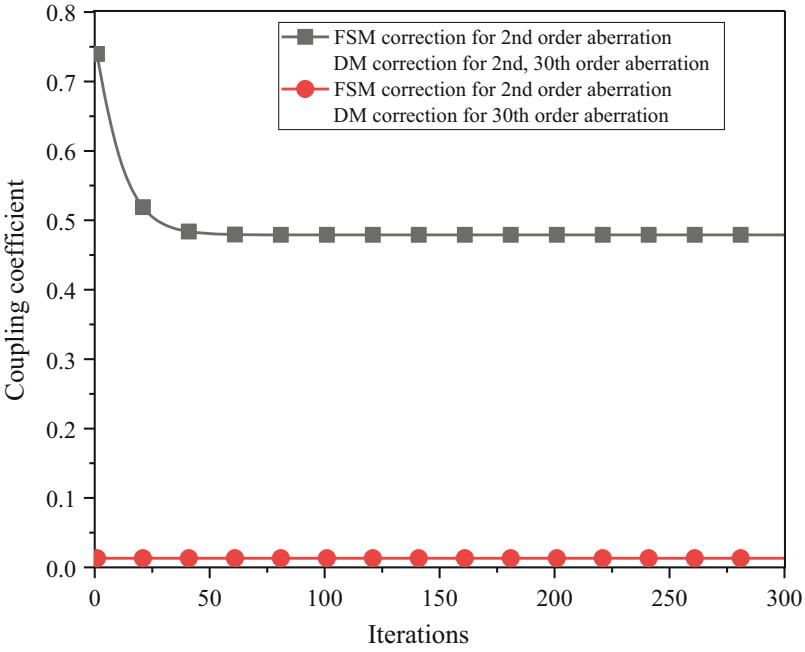


Fig. 25.4 Coupling-coefficient curves between FSM and DM

beam was incident on the surface of the FSM and then on the surface of the DM. The FSM and DM corrected the specified aberrations as needed. The beam corrected by the FSM and DM was incident onto a 4f system composed of lenses L_3 and L_4 for beam reduction.

The wavefront sensor detected the focal point of lens L_4 and transmitted the detected wavefront information to a wavefront controller (computer). The computer calculated the control voltage using an appropriate control algorithm, and sent it to the FSM and DM for closed-loop correction of the distorted wavefront over multiple iterations. This formed a typical adaptive-optics dual-deformable-mirror closed-loop correction system.

According to the experimental diagram shown in Fig. 25.5, the experimental parameters of the PI controller were $k_{i_w} = 0.999$, $k_{p_w} = 0.1$, $k_{i_t} = 0.999$, and $k_{p_t} = 0.1$. The decoupling algorithm studied in this article was used for indoor experiments. Figure 25.6 shows the coupling-coefficient curve between the FSM and DM.

Figure 25.6a presents the coupling-coefficient curve when the FSM corrects only the first- and second-order aberrations, and the DM corrects only the third-order aberrations. The approximate trend of the curve indicates that the coupling coefficient is approximately 1.117×10^{-3}. Figure 25.6b displays the coupling-coefficient curve when the FSM only corrects the first- and second-order aberrations, and the DM only corrects the tenth-order aberrations, with a coupling

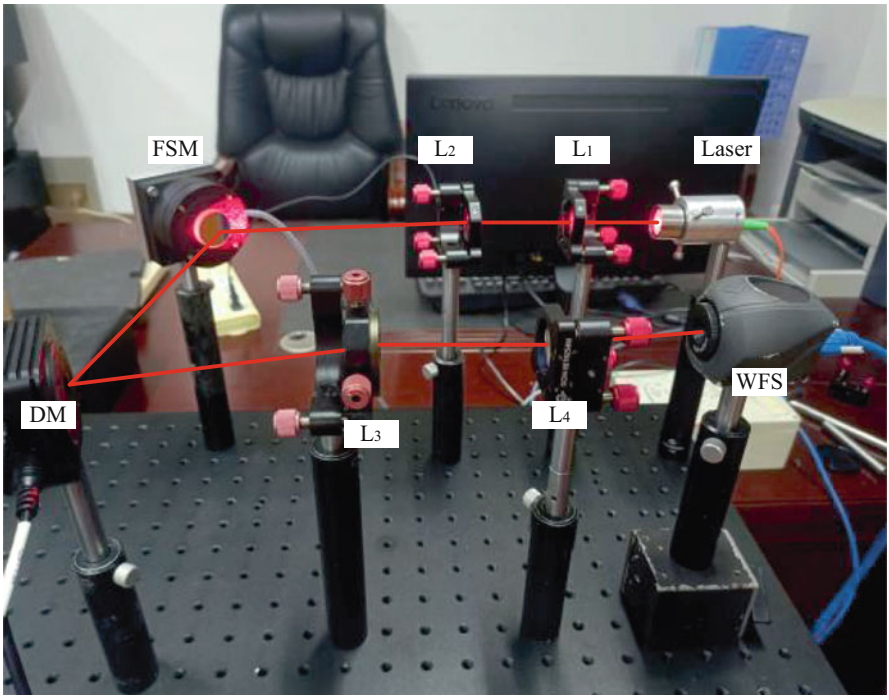


Fig. 25.5 Experimental diagram of a typical adaptive-optics dual-deformable-mirror wavefront-correction system

coefficient of approximately 1.824×10^{-2}. Figure 25.6c shows the coupling-coefficient curve when the FSM only corrects the first- and second-order aberrations, and the DM only corrects the 25th-order aberrations, with a coupling coefficient of approximately 7.81×10^{-3}.

Experiments demonstrated that the decoupling algorithm was effective in suppressing the coupling effect between the FSM and the DM. It performed graded correction of aberrations, selectively allocated any order of aberrations for correction, and independently corrected any order of aberrations. The coupling coefficients reached levels of 10^{-2} and below, effectively suppressing the coupling effect between the FSM and DM.

Coupling Analysis

The coupling coefficient is calculated from the directional cosine of the woofer and tweeter phase compensations. This coefficient represents the degree of correlation between them [31]. Through research, the coupling effect of the dual-deformable-mirror wavefront-correction system can be effectively reduced, independently correcting each order of aberration and any order of aberrations, as well as correcting

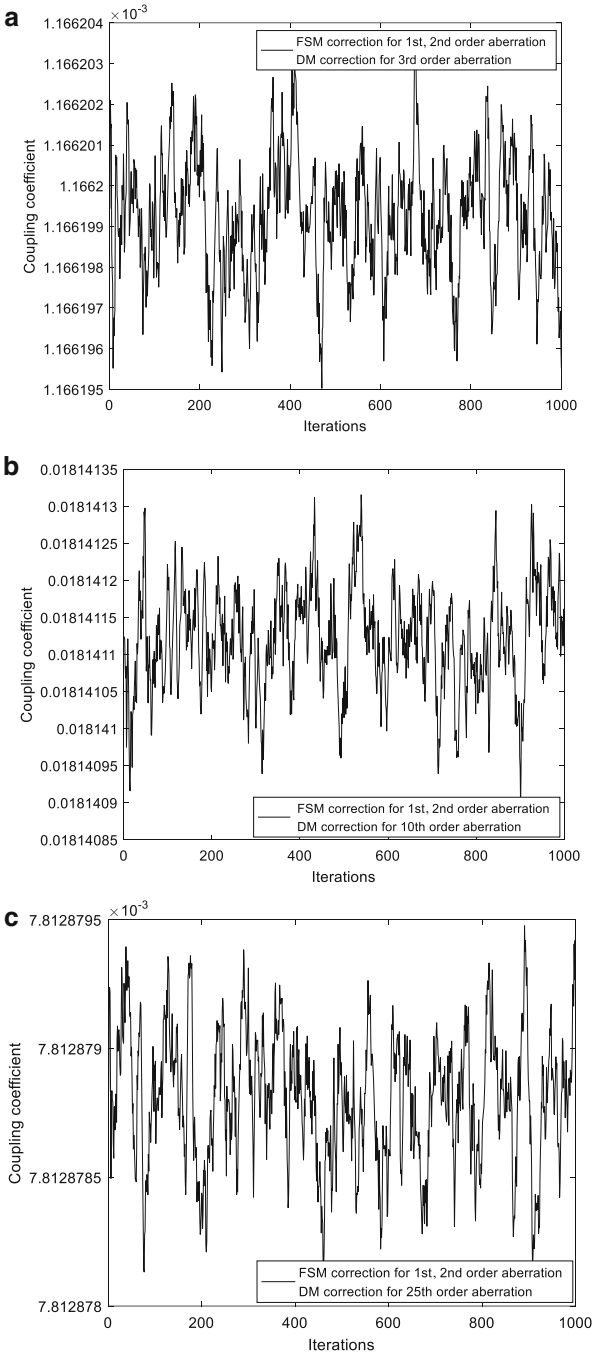


Fig. 25.6 Coupling-coefficient curves between FSM and DM

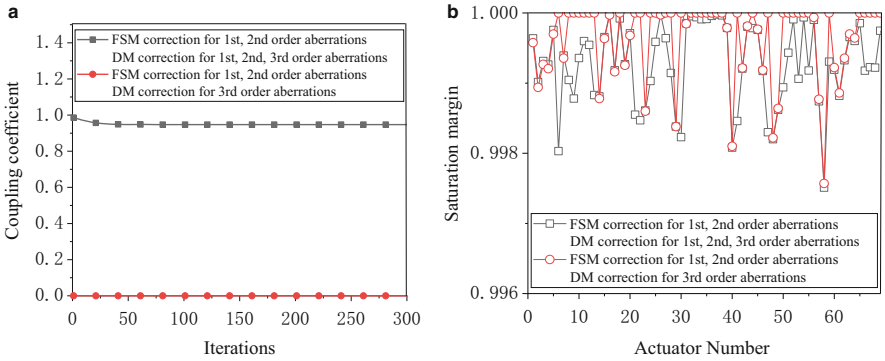


Fig. 25.7 (**a**) Coupling-coefficient curve. (**b**) Saturation margins of DM actuators

specific targeted aberrations in different scenes. This makes the dual-deformable-mirror wavefront-correction system more widely applicable and lays the foundation for its application in large-aperture telescopes.

For example, when correcting third-order aberrations, the mode coefficients of the aberrations are 0.657761 μm, 2.42713 μm, and 0.0423 μm. Figure 25.7 shows the saturation margin of the DM actuator and the coupling-coefficient curve between the FSM and DM.

Figure 25.7a shows the coupling-coefficient curve between the FSM and DM. Figure 25.7b shows the saturation margin of the DM actuator. The black line in Fig. 25.7a represents the coupling-coefficient curve when correcting the first- and second-order aberrations of the FSM and the first-, second-, and third-order aberrations of the DM. The red line represents the coupling-coefficient curve when correcting the first- and second-order aberrations of the FSM and the third-order aberration of the DM.

The black line in Fig. 25.7b represents the saturation margin of the DM actuator, which corresponds to the black line shown in Fig. 25.7a. The red line represents the saturation margin of the DM actuator, which corresponds to the red line in Fig. 25.7a. As indicated by the black line in Fig. 25.7b, the saturation margin of some of the DM actuators decreases rapidly, which reduces the ability to correct aberrations and affects the correction of higher-order aberrations. This seriously affects the stroke of the DM actuator, resulting in wasted resources and increased costs. Therefore, a decoupling study of the dual-deformable-mirror wavefront-correction system is crucial.

Conclusion

In this study, a decoupling algorithm was proposed for a dual-deformable-mirror wavefront-correction system. The algorithm utilized Zernike coefficients to directly represent the compensating ability of the deformable mirror and enabled the

independent correction of each order of aberration. Based on the expression for the coupling coefficient in a typical dual-deformable-mirror wavefront-correction system, this study analyzed the coupling effect of correcting different aberrations and drew the following conclusions:

1. A typical dual-deformable-mirror wavefront-correction system can independently correct various orders of aberrations as well as specific targeted aberrations in different scenes, thereby effectively suppressing the coupling effect between deformable mirrors.
2. The algorithm proposed in this study is simple and feasible, and was verified to be applicable to a dual-deformable-mirror wavefront-correction system with a wavefront sensor. This study has great practical significance in engineering and provides an important reference for practical applications.

References

1. X. Ke. Adaptive Optics Theory and Its Application in Optical Wireless Communication [M]. Beijing: Science Press, 2021.
2. Liu W, Dong L, Yang P, et al. Zonal decoupling algorithm for dual deformable mirror adaptive optics system [J]. Chinese Optics Letters, 2016, 14(2): 020101–020105.
3. Jiang W, Li H. Hartmann-Shack wavefront sensing and wavefront control algorithm [J]. SPIE, 1990, 1271(4): 82–93.
4. Vorontsov M A. Decoupled stochastic parallel gradient descent optimization for adaptive optics: integrated approach for wave-front sensor information fusion [J]. JOSA A, 2002, 19(2): 356–368.
5. Hu S J, Xv B, Wu J, et al. Research on Aberration Decoupling of Dual Deformable Mirror Adaptive Optical System [J]. Acta Optica Sinica,2005(12):1687–1692.
6. Hu S J. Performance of an adaptive optics system with dual deformable mirrors [J]. Journal of Electronic Science and Technology, 2009, 7(3): 277–280.
7. Conan R, Bradley C, Hampton P, et al. Distributed modal command for a two-deformable-mirror adaptive optics system [J]. Applied Optics, 2007, 46(20): 4329–4340.
8. Zou W, Qi X, Burns S A. Wavefront-aberration sorting and correction for a dual-deformable-mirror adaptive-optics system [J]. Optics Letters, 2008, 33(22): 2602–2604.
9. Ye H w, Li X y. Joint closed-loop control method for multiple deformable mirrors in adaptive optical systems [J]. China Laser, 2009, 36(s2):0258–7025
10. Sredar N, Ivers K M, et al. A correction algorithm to simultaneously control dual deformable mirrors in a woofer-tweeter adaptive optics system [J]. Optics express, 2010, 18(16): 16671–16684.
11. Lei X, Wang S, Yan H, et al. Double-deformable-mirror adaptive optics system for laser beam cleanup using blind optimization [J]. Optics express, 2012, 20(20): 22143–22157.
12. Liu W, Dong L, Yang P, et al. A Zernike mode decomposition decoupling control algorithm for dual deformable mirrors adaptive optics system [J]. Optics Express, 2013, 21(20): 23885–23895.
13. Li Z, Cao J, Zhao X, et al. Combinational-deformable-mirror adaptive optics system for atmospheric compensation in free space communication [J]. Optics communications, 2014, 320(27): 162–168.
14. Zhou R, Shen F, Ye H, et al. Double deformable mirrors' control based on voltage decoupling for adaptive optics system [J]. SPIE, 2015, 9255(3): 799–804.

15. Cheng T, Liu W, Pang B, et al. A slope-based decoupling algorithm to simultaneously control dual deformable mirrors in a woofer–tweeter adaptive optics system [J]. Chinese Physics B, 2018, 27(7): 1–9.
16. Cheng T, Liu W J, Yang K J, et al. Testing for a slope-based decoupling algorithm in a woofer-tweeter adaptive optics system [J]. Applied Optics, 2018, 57(13): 3357–3364.
17. Cheng T, Liu W J, Yang K J, et al. Projection-based decoupling algorithm for a woofer–tweeter adaptive optics system [J]. Optical Engineering, 2018, 57(6): 1–10.
18. Cheng T, Xu Z X, Yang K, et al. A decoupling control algorithm for Woofer–Tweeter adaptive optics system in generalized irregular pupil region [J]. Optics Communications, 2020, 472(4): 1–12.
19. Kong L, Cheng T, Yang P, et al. Decoupling control algorithm based on numerical orthogonal polynomials for a woofer-tweeter adaptive optics system [J]. Optics Express, 2021, 29(14): 22331–22344.
20. Yang S, Ke X. Experimental study on adaptive optical wavefront correction with dual mirrors in free space optical communication [J]. Optik, 2021, 242(5): 167146.
21. Wang Y, Zhang X, Li D, et al. A Global Decoupling Control Algorithm for LC-DM Adaptive Optics System [J]. IEEE Access, 2021, 9(10): 142302–142313.
22. Wang Y, Li D, Jin C. Eigenmode Wavefront Decoupling Algorithm for LC–DM Adaptive Optics Systems [J]. Applied Sciences, 2022, 12(15): 1–16.
23. W. Liu, B. Xv, W. Sun, Decoupling Control Algorithm Used for Wave Front Sensor‐Free Adaptive Optics System with Dual Deformable Mirrors Based on Mode Project‐Constraint [J]. Chinese Journal of Lasers, 2023,50(06), 106–112.
24. Zou W, Qi X, Burns S A. Woofer-tweeter adaptive optics scanning laser ophthalmoscopic imaging based on Lagrange-multiplier damped least-squares algorithm [J]. Biomedical Optics Express, 2011, 2(7): 1986–2004.
25. M. Hu. A Review of Simulation Methods for the Optical Wavefront Distorted by Atmospheric Turbulence [J]. Optics & Optoelectronic Technology, 2003,1(05), 10–12.
26. J. Liang, S. Lin, M. Han, et al., Research progress of mode method in adaptive optics [J]. Modern Applied Physics 2023,2(14), 1–16.
27. Noll R J. Zernike polynomials and atmospheric turbulence [J]. Journal of the OpticalSociety of America, 1976, 66(03): 207–211.
28. Z. Hu, W. Jiang, Simulation of the Optical Wavefront Distorted by Atmospheric Turbulence [J]. Opto-Electronic Engineering,1995, 22(02), 50–56.
29. Ch. Xv, Sh. Hao, Y. Wang, et al., Atmospheric Turbulence Distortion Wave-Front Reconstruction Based on Modal Mutual Information [J]. Journal of Air Force Engineering University,2020, 21(02), 65–70.
30. Ke X, Wu J, Hao J. Distorted wavefront reconstruction based on compressed sensing [J]. Applied Physics B, 2022, 128(6): 107–117.
31. J. Wu. The Beauty of Mathematics in Computer Science [M]. Beijing: Posts and Telecommunications Press, 2012.

Optical Phase-Locked Loops

26

Contents

Introduction 942
Development of Phase-Locked Loops 942
Development of Electric Phase-Locked Loops 942
Development of Optical Phase-Locked Loops 943
Electronic Phase-Locked Loops 951
Basic Structure of an Electronic Phase-Locked Loop 951
Classification of Electric Phase-Locked Loops 952
Parameters of Electronic Phase-Locked Loops 961
Applications of Electric Phase-Locked Loops 963
Optical Phase-Locked Loops in Optical Communication 964
Basic Structure of an Optical Phase-Locked Loop 964
Classification of Optical Phase-Locked Loops 968
Parameters of the Optical Phase-Locked Loop 974
Performance Comparison of Various Optical Phase-Locked Loops 975
Optical Phase-Locked Loops in Optical Wireless Coherent Communication 977
Optical Phase-Locked Loop for Coherent Optical Communication 977
Typical Applications of Optical Phase-Locked Loops 978
Summary 980
References 980

Abstract

Phase-locked loop (PLL) plays an important role in wireless communication navigation, radar, and other fields. It affects the data processing rate of the whole system and the extraction of clock phase information. The optical phase-locked loop (OPLL) has the advantages of high-speed real-time reception, low power consumption, and strong ability to resist background light interference and Doppler frequency shift, and is the key circuit module of coherent optical communication system. Its performance directly affects the performance of optical communication system. The research of optical phase-locked loop is mainly aimed at optimizing and improving its parameters. This chapter reviews

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_26

the research progress of the technology from electric phase-locked loop to optical phase-locked loop. Based on an introduction to the principle and key components, the classification and parameters of the electric phase-locked loop and optical phase-locked loop are introduced in detail, and the performance of various optical phase-locked loops is compared. This chapter summarizes the optical phase-locked loop technology in wireless coherent optical communication systems, and finally prospects its future development direction, which can provide reference for the research and development of phase-locked loop technology in different fields.

Keywords

Wireless communication · Optical communication · Coherent optical communication · Phase-locked loop

Introduction

Traditional phase-locked loop technology was first applied in receiving systems for wireless communication. The clock phase is extracted from the received signal, and the local clock and the received-signal clock are synchronized through feedback control [1]. Optical phase-locked loop (OPLL) technology originated from coherent optical communication systems [2, 3]. Compared with the intensity modulation/direct detection (IM/DD) communication mode in an optical communication system, the sensitivity of coherent communication can be improved by 20 dB to 23 dB.

Coherent detection is the best solution for achieving high system sensitivity and suppressing out-of-band noise [4]. Coherent optical communication can be divided into homodyne and heterodyne coherent reception, depending on whether the frequencies of the signal light and local light are consistent. The sensitivity of homodyne coherent reception is 3 dB higher than that of heterodyne reception [5].

In homodyne coherent reception, the phases of the signal light and the local oscillator light should be strictly synchronized, which requires the communication system to be able to control the laser frequency. As a closed-loop optical-band control system [6], an optical phase-locked loop can match the phases of two beams of light; thus, an optical phase-locked loop becomes the core technology of a coherent receiver.

Development of Phase-Locked Loops

Development of Electric Phase-Locked Loops

In 1665, Huygens [7] proposed the concept of a synchronous oscillator by observing the spontaneous weak coupling of a pendulum. This was the first concept of phase-locking. Limited by the level of scientific and technological development at that time, the academic community did not focus on the concept of phase-locking.

In 1919, Vincent and Appleton [8] discussed the phase-locking concept in the literature and proposed a method to synchronize oscillators. In the 1930s, radio technology was in the early development stages and some mathematical theories of phase-locked loops appeared [9]. In 1932, French De Bellescize [10] published the theory of synchronous detection; then, he provided a mathematical description of a phase-locked loop, and applied it to synchronous detection.

In 1940, researchers first applied phase-locked loops (PLL) to television sets [11]; they effectively suppressed the screen jitter caused by external noise interference, improved image-synchronization performance, and made TV pictures clearer. By 1947, phase-locked loop technology was widely used in TV horizontal synchronous circuits [12].

In the 1950s, Jaffe and Rechtin [13] applied phase-locked loop technology in a missile-tracking filter device and theoretically analyzed the phase-locked loop, including the noise effect. The Soviet Union launched the first artificial earth satellite in 1957, and phase-locked loop technology was first applied in this field, opening the human space era [14]. In 1965, an analog-integrated phase-locked loop [15] was developed using analog-circuit technology, which greatly reduced the cost of phase-locked loop circuits.

In the 1970s, phase-locked loops appeared in digital–analog hybrid systems [16]. Some circuits in the loop structure were digitized, whereas the rest were still composed of analog circuits. The charge-pump phase-locked loop (CPPLL) proposed by Gardner [17] in 1980 is a typical digital–analog hybrid phase-locked loop. In 1988, Johnson et al. [18] designed a phase-locked loop structure based on a CMOS voltage-control delay, which dynamically adjusted the clock period, improved the loop stability, and enhanced the anti-noise performance of the system as a whole.

In 1993, Riley [19] proposed a fractional frequency-division phase-locked loop structure to improve the accuracy of the output frequency. This fractional frequency division was achieved using a Δ-Σ accumulator modulator. With the development of digital circuits, in 2003, Staszewski [16] proposed the structure of an all-digital phase-locked loop (ADPLL). In the twenty-first century, with the development of electronic communication, the structure of phase-locked loops has improved from single loops to composite loops, and has been developing toward low power consumption, high frequency, and wide range.

Development of Optical Phase-Locked Loops

Since Enlog et al. [20] first used a He–Ne laser to perform laser phase-locked experiments in 1964, many structural implementation schemes have been developed for optical phase-locked loops; for example, researchers have successfully conducted homodyne and heterodyne phase-locked laser experiments, and phase-locked experiments with distributed Bragg reflection lasers.

In 1977, Mengali et al. [21] analyzed the performance of an optical phase-locked loop driven by a photodetector current in an optical receiver and analyzed the loop-tracking characteristics of pulse-amplitude modulation and sine-wave modulation.

In 1983, Steele [22] designed a heterodyne optical phase-locked loop structure based on GaAlAs semiconductor lasers. The experimental results showed that the locking range of the optical phase-locked loop reached 40 MHz.

From 1985 to 1986, Kazovsky [23, 24] researched various types of optical phase-locked loops, including heterodyne decision-driven, homodyne balanced-detection, and homodyne decision-driven phase-locked loops.

In 1989, Germany launched a solid-state laser communication system (SOLACOS) [25] that used the Costas optical phase-locked loop as the terminal model to realize one-way communication between low-orbit satellites and synchronous-orbit satellites.

In 1990, the European Space Agency (ESA) demonstrated simulated space experiments using a coherent-reception terminal with the Costas loop, based on binary phase-shift keying (BPSK) modulation with a transmission rate of 140 Mbit/s [26].

In 1990, Kazovsky et al. [27] conducted a coherent-reception experiment on a balanced optical phase-locked loop using a Nd:YAG laser. The sensitivity of the receiver was 25 photons/bit when the phase-shift keying homodyne transmission rate was 140 Mbit/s, and 332 photons/bit when the transmission rate was 2 Gbit/s. In the same year, Kahn et al. [28] used an external-cavity semiconductor laser to conduct an experiment on a BPSK homodyne balanced phase-locked loop system. The receiver sensitivity was 83 photons/bit when the transmission rate was 4 Gbit/s.

Schopflin et al. [29] conducted an experimental coherent-reception study based on a Costas phase-locked loop. The experimental results showed that the sensitivity could reach 11 photons/bit at a communication rate of 140 Mbit/s and 21 photons/bit at 565 Mbit/s. Norimatus et al. [30, 31] carried out an experiment based on a Costas phase-locked loop receiver. When the transmission rate was 5 Gbit/s, the measured sensitivity was 93 photons/bit. They then carried out an experimental study on a decision-driven phase-locked loop. When the transmission rate was 10 Gbit/s, the corresponding sensitivity was 297 photons/bit.

Before modulating the data signal, Wandemoth [32] added a zero-bit signal as the synchronization bit every eight bits, thus improving the structure of the optical phase-locked loop and obtaining a synchronization-bit phase-locked loop. The sensitivity measured in the experiment was 20 photons/bit when the communication rate was 565 Mbit/s. As shown in Table 26.1, a relationship between the communication bandwidth and coherent-detection sensitivity has been experimentally proven.

In 1991, Bopp et al. [33] used a receiver based on a Nd:YAG solid laser to conduct a communication experiment using BPSK zero-difference detection and differential phase-shift keying (DPSK) modulation based on a Costas phase-locked loop. The communication rate reached 600 Mbit/s.

In 1992, Hornbachner et al. [34] used a Costas phase-locked loop to conduct a phase-synchronization experiment between the input signal and the local optical signal. Compared to the limitation of ideal shot noise, the sensitivity of the receiver was improved by 5.9 dB.

In 1995 and 1996, Wang et al. [35, 36] successively carried out long-distance communication experiments based on homodyne balanced and homodyne Costas

Table 26.1 Various homodyne coherent-detection experiments based on optical phase-locked loops

Year	Author	Type of optical phase-locked loop	Communication rate	Detection sensitivity (photons/bit)
1990	Kazovsky	Balanced loop	140 Mbit/s 2 Gbit/s	25 332
1990	Kahn	Balanced loop	4 Gbit/s	83
1990	Schopflin	Costas loop	140 Mbit/s 565 Mbit/s	11 21
1990	Norimatus	Costas loop Decision-driven loop	5 Gbit/s 10 Gbit/s	93 297
1991	Wandemoth	SyncBit loop	565 Mbit/s	20
1992	Hornbachner	Costas loop	140 Mbit/s	28
2005	Herzog	Third-order dither loop	400 Mbit/s	36

loops, and theoretically analyzed the performance of an optical amplifier in an optical phase-locked loop system under the effect of noise.

In 1996, the European Space Agency implemented a new generation of inter-satellite coherent optical communication projects called short-range optical inter-satellite links (SROIL) [37]. The receiver used BPSK homodyne coherent-detection technology as the modulation mode. When the transmission distance was 4000 km, the transmission rate of the receiver was 1.5 Gbit/s, and the measured bit-error rate was approximately 10^{-9}. The project also carried out medium-long distance inter-satellite laser communication experiments, and the receiver used BPSK homodyne coherence technology. When the transmission distance was 6000 km, the measured transmission rate was 6.5 Gbit/s.

In 1999, Djordjevic [38] studied an optical heterodyne system based on phase-shift keying (PSK) and a Costas phase-locked loop for multichannel frequency-division multiplexing. In 1999, Kudielka and Klaus [39] designed an optical phase-locked loop receiver that combined the advantages of synchronous bit and residual carriers. The receiver did not require complex electronic equipment or a special modulation circuit for the transmitter.

In 2002, the Deutsche Zentrum für Luft- und Raumfahrt (DLR) and TESAT jointly launched a laser communication terminal for the SAR-X (LCTSX) project. This project used a Costas optical phase-locked loop homodyne coherent-optical communication system with a laser wavelength of 1064 nm [40]. The error rate of the system was approximately 10^{-9}.

In 2004, Camatel et al. [41] proposed a subcarrier optical phase-locked loop to improve a decision-driven phase-locked loop, which was implemented based on external modulation.

In 2005, Smutny et al. [42] investigated Costas phase-locked loop receiving technology with BPSK modulation on a communication optical terminal. They achieved a transmission rate of 6.75 Gbit/s, with a 6000-km transmission distance, and a measured bit-error rate of 10^{-8}.

In 2005, Herzog et al. [43] investigated a third-order oscillatory optical phase-locked loop supported by the ESA project. The experimental results showed that the communication rate was 400 Mbit/s and the reception sensitivity reached 36 photons/bit.

In 2005, Zhang et al. [44] introduced a control system to correct the laser phase in optical phase locking, locked the phases between the independent elements of the array laser, and effectively suppressed the laser-phase noise.

In 2006, Camatel et al. [45] used a subcarrier optical phase-locked loop to detect the zero-difference coherence of amplitude shift keying (ASK) and phase-shift keying (PSK) modulated signals. They tested the coherent communication between ASK and PSK at a transmission rate of 2.5 Gbit/s using a 10-Gbit/s PSK transmission system.

In 2006, the German TESAT Research Center launched a plan to establish a low-earth orbit (LEO) to geosynchronous earth orbit (GEO) inter-satellite laser link with a transmission rate of 2 Gbit/s using BPSK-modulation homodyne coherent-detection technology. The link was further extended to GEO–GEO and GEO–ground laser links [46, 47], which had transmissions rate of up to 2 Gbit/s.

In 2006, the National Intelligence and Communications Institute of Technology (NICT) of Japan developed a real-time coherent receiver model [48]. The front end of the receiver included a 90° optical hybrid, two tunable lasers, and two balanced optical detectors. The system demodulated coherent optical signals at a rate of 3 Gbit/s.

In 2007, Yariv et al. [49] proposed a new control system consisting of a multistage optical phase-locked loop to control the beam frequency and phase, and locked the phases of two 1064-nm semiconductor lasers.

In 2007, the German TerraSAR-x satellite LCTSX orbited the Earth at a polar orbit of 514.8 km. LCTSX adopted an optical coherent receiver based on a homodyne BPSK optical phase-locked loop [50] with a data-transmission rate of 5.5 Gbit/s.

In 2007, Zhang et al. [51] proposed a new optical phase-locked loop mechanism for clock extraction. It was based on the four-wave mixing effect of an electric absorption modulator (EAM). It was experimentally determined that the system could extract the clock information of a 40/80 Gbit/s optical time-division multiplexing optical-data signal.

In 2007, Quan et al. [52] established a new type of optical phase-locked loop, which was based on the four-wave mixing effect of a voltage-controlled optical pulse source and an optical fiber; the extinction ratio of the phase detector was greater than 30 dB.

In 2008, Camatel et al. [53] built a subcarrier optical phase-locked loop homodyne receiver based on BPSK modulation, and demonstrated a nonlinear-subcarrier optical phase-locked loop based on a decision-drive scheme. The performance of the decision-driven receiver was 3 dB better than that of the pilot-carrier receiver based on 8B/10B coding.

In 2008, the United States' Near Field Infrared Experience (NFIRE) satellite and Germany's TerraSAR-X satellite completed a LEO–LEO inter-satellite communication test. The system had a communication rate of 5.625 Gbit/s and used a BPSK-

modulation Costas phase-locked loop homodyne-detection structure [50]. Since then, the NFIRE satellite of the United States conducted a satellite–ground laser-communication link with a European ground station, with a communication rate of 5.625 Gbit/s, which was the first satellite–ground coherent-laser link test in the world [54].

In 2009, Ristic et al. [55] successfully integrated the core of an optical phase-locked loop using photonic integrated-circuit technology. In 2009, optical inter-satellite communication based on the TESAT laser communication terminal (LCT) in Europe operated on a LEO satellite for more than 2 years [56]. The communication terminal used a modulation scheme based on homodyne BPSK and achieved a bit-error rate of 10^{-9} at a full-duplex data rate of 5.625 Gbit/s. LCT not only shows its characteristics in LEO–LEO inter-satellite links but can also be used in LEO–ground link communication to study the propagation of beams in the atmosphere.

In 2009, Xu et al. [57] proposed a balanced homodyne receiver using Costas-loop feedback to detect weak coherent states and used an experimental BPSK-modulation device to conduct an optical carrier phase-tracking experiment.

In 2009, Wang [58] studied a coherent-optical communication homodyne BPSK system, completed an efficient optical homodyne detection and communication signal-demodulation experiment, and analyzed the noise performance of balanced and decision-feedback optical phase-locked loops. To reduce the phase difference between the signal light and the local oscillator light caused by phase noise, the optical phase-locked loop needs a larger bandwidth, but the larger the loop bandwidth, the greater the noise. In addition, shot noise will bring additional phase-locked errors.

In 2010, Steed et al. [59] studied a tunable coherent terahertz source based on a hybrid-integrated optical phase-locked loop. The system had the potential to generate signals up to 1.8 THz, the linewidth of the heterodyne signal between the lasers was less than 1 kHz, and the offset frequency was less than −80 dBc/Hz when the phase noise was at 10 kHz.

In 2010, Yariv et al. [60] researched composite phase-locked technology based on distributed-feedback (DFB) lasers, using optical phase modulators to expand the bandwidth of the loop and reduce residual phase errors. This experiment improved the coherence between master and slave lasers in an optical phase-locked loop.

In 2010, Sakamoto et al. [61] proposed a new structure for a digital Costas optical phase-locked loop, based on a subcarrier optical phase-locked loop, to realize the real-time homodyne demodulation of BPSK signals with a transmission rate of 40 Gbit/s.

In 2010, Du [62] analyzed optical phase-locked loop technology and backward phase-synchronization and correction technology, based on frequency-offset extraction, and adopted this technology to keep the quadrature-in-phase branch signal strictly orthogonal, thus eliminating phase noise at the transmitter.

In 2011, the NICT of Japan designed optical-injection phase-locked loop technology. The phase-locked loop used the structure of the Costas phase-locked loop [63]. The experiment showed that when the signal laser frequency changed more

than 9 GHz at a frequency-deviation rate of up to 1 Hz, the 10 Gbit/s BPSK link successfully maintained a stable frequency and a phase-locked state, meeting the requirements of LEO-to-optical-ground-station coherent links.

In 2011, the Japan Aerospace Exploration Agency (JAXA) [64] designed a homodyne coherent receiving terminal based on BPSK modulation, which was applied in a laser-link experiment between GEO and LEO satellites in a geosynchronous orbit. The transmission rate was up to 2.5 Gbit/s, and the demodulation sensitivity was measured to be −49.1 dBm when the initial frequency offset was ±7.5 GHz.

In 2011, Steed et al. [65] conducted a heterodyne optical phase-locked loop experiment. They used a custom-designed feedback circuit to lock the phase noise of a heterodyne signal between two lasers with a 1553-nm emission wavelength and 1.1-MHz linewidth below −90 dBc/Hz and a frequency offset greater than 20 kHz. The variance of the phase error was 0.04 rad^2 in a 10-GHz bandwidth.

In 2011, Qiu [66] carried out a theoretical analysis and mathematical modeling experiments on a Costas optical phase-locked loop and studied the loop-stability problem. The results showed that with an increase in the loop-damping coefficient, the loop-acquisition speed would be accelerated. When the damping coefficient was 0.64, the maximum value of the normalized oscillation frequency was 0.74. In this case, the loop stability was the best.

In 2011, Ding [67] designed an optical phase-locked loop system based on a ring laser. Through simulation analysis, it was found that when the total gain of the loop was 2.7 MHz, the system could lock the phase for a short time. The locking time was generally a few seconds, with a minimum of approximately 1 s.

In 2012, Park et al. [68] established a highly integrated 40-Gbit/s coherent-optical receiver based on a Costas phase-locked loop structure and experimentally achieved a 1.1-GHz loop bandwidth and 120-ps loop delay. In 2012, Lu et al. [69] first proposed a highly integrated optical phase-locked loop with a phase/frequency detector and a single sideband mixer and measured its loop bandwidth up to 400 MHz.

In 2012, Mizutori et al. [70] implemented a 12.5-Gbit/s BPSK zero-difference decision-drive loop experiment with a 3-kHz linewidth outer-cavity semiconductor laser. The performance of this phase-locked loop enabled the sensitivity of the receiver to match that of a digital coherent receiver.

In 2012, Xu [71] designed a fully digital optical phase-locked system with a feedback control circuit for an atomic interferometer using an external-cavity semiconductor laser. The phase-locked loop achieved a loop bandwidth of 570 kHz and a residual phase variance of 0.14 rad^2.

In 2012, Peng et al. [72] used an optical phase-locking method to track the frequency of two semiconductor DFB lasers with a 1550-nm wavelength. The frequency difference between the master and slave lasers was constant at 11 GHz.

In 2013, JAXA studied a homodyne coherent-optical receiving system based on a BPSK Costas ring. The receiver also had good stability in the presence of a ±7.5-GHz doppler-frequency shift and successfully established a GEO–LEO data link

with a transmission rate of 7.2 Gbit/s [73]. The measured bit-error rate was 10^{-6} and the sensitivity was −43 dBm.

In 2013, Ding et al. [74] studied an optical phase-locked loop device for locking the frequency of coherent lasers. In the experiment, when the beat signal was 6.834 GHz, it had a bandwidth of −3 dB at 1 Hz.

In 2013, Meng et al. [75] designed an optical phase-locked loop using a digital phase-locked loop chip with a frequency-locking bandwidth of 1–7 GHz. In the experiment, the magnitude of the beat-signal noise was reduced from megahertz to hertz.

In 2014, Balakier et al. [76] investigated the properties and performance of an integrated balanced optical phase-locked loop used to generate tunable millimeter/THz waves, such as tunability and long-term frequency locking.

In 2014, Shi et al. [77] conducted an experiment with a Costas optical phase-locked loop with a transmission rate of 5 Gbit/s BPSK and measured a sensitivity of −41.4 dBm when the bit-error rate was 10^{-9}.

In 2014, Yue [78] conducted an experiment on a homodyne coherent-optical communication system with a balanced phase-locked loop. When the communication rate was 1.25 Gbit/s, the measured communication sensitivity was −51 dBm and the error rate was 9.72×10^{-5}.

In 2015, Koga et al. [79] studied a Costas-ring homodyne receiver with a 20-Gbit/s communication rate and quadrature phase-shift keying (QPSK) modulation mode, based on the concept of an optical synchronization network.

In 2015, Zhang et al. [80] designed a digital optical phase-locked system based on a Costas loop. When the frequency-difference range was within 140 MHz, the lock-in time could reach 1 h. A tracking experiment was conducted with a 20-MHz/s Doppler frequency shift.

In 2015, Jiang [81] proposed a heterodyne optical phase-locked loop based on the internal modulation of ultra-narrow linewidth lasers, and a phase-locking experiment of 100-μW strong light and 10-nW weak light for 10 h was successfully conducted.

In 2016, Wang et al. [82] designed an optical phase-locked loop with a frequency discrimination and control subsystem. The experimental results showed that the optical phase-locked loop locked the phase within 0.2 ms, and accurately tracked the sinusoidal phase jitter with an amplitude of $\pm\pi/2$ at a 1-kHz jitter frequency.

In 2016, Ma et al. [83] designed an optical phase-locked loop circuit based on an ADF41020 phase-locked loop chip to lock two laser beams with frequency differences ranging from 4 to 18 GHz.

In 2017, Wang et al. [84] demonstrated a method based on a balanced phase-locked loop for fast laser-frequency acquisition and conducted a phase-locked experiment with a large initial-frequency difference. When the frequency acquisition range was 164 MHz, the measured acquisition and locking time was 440 ms.

In 2017, Chang et al. [85] conducted research on optical phase-locked loop technology for inter-satellite high-speed coherent laser communication. When using BPSK for modulation, when the transmission rate was 5 Gbit/s, the measured sensitivity reached −41.4 dBm, and the receiver error rate was 10^{-7}.

In 2017, Zhou et al. [86] designed a coherent receiver based on a Costas optical locking ring with an auxiliary acquisition and locking circuit. It compensated for the Doppler shift caused by the relative motion of satellites in real time. The sensitivity of the receiver was -39.5 dBm and the bit-error rate was 10^{-9} when the communication rate was 10 Gbit/s. In 2018, they [87] further proposed the structure of a Costas covariance coherent receiver with 90° mixed signals in space optics. The results showed that the sensitivity of the Costas coherent receiver with an optimal signal power-distribution ratio was close to the ScatterNet noise limit.

In 2018, Ashtiani et al. [88] studied a high-resolution optical-frequency synthesizer with an integrated electro-optical phase-locked loop. It could generate optical signals with a subhertz-tuned resolution of 192 THz, within the range of 5 THz, with a coarse-tuned resolution of 20 MHz and a fine-tuned resolution of 1 Hz.

In 2018, Wang et al. [89] added an adjustable advanced phase-shifting circuit and adjustable-gain unit into the single-loop feedback loop of the optical phase-locked loop. They determined that the residual phase noise of the optical phase-locked loop was as low as 0.012 rad^2, which can reduce the laser linewidth from 2 MHz to 10 kHz.

In 2018, Liu et al. [90] simulated the phase-locking process of an optical phase-locked loop by building a BPSK zero-difference coherent-optical communication system model. The results showed that the Costas optical phase-locked loop could lock the phase of the signal light to approximately 1 μs when the frequency difference was less than 50 MHz.

In 2018, Yue et al. [91] adopted an improved Costas optical phase-locked loop and a digital control algorithm to suppress the phase error and track the Doppler shift to realize zero-error detection. The experimental results showed that the receiver-sensitivity loss caused by the phase error was approximately 1 dB. The receiver could track the Doppler shift at 100 MHz/s.

In 2019, Tao [92] established a pulse-laser balanced phase-locked loop model, based on broadband-laser phase-locked technology, and analyzed that the loop-capture frequency band could reach 120 MHz. They then built a loop experimental platform and measured the loop delay as only 1.2 ns; the loop bandwidth reached 40 MHz, thus, realizing laser phase locking.

In 2019, Liu [93] proposed a Costas phase-locked loop broadband-tuning method based on an acousto-optic frequency shifter. In the 10-GHz capture range, the loop bandwidth was 1.5 MHz, and 1×10^{-8} was the measurement standard. When the communication rates were 10 Gbit/s and 5 Gbit/s, the measured sensitivities were -38.7 dBm and -41.4 dBm, respectively.

In 2019, Pan et al. [94] conducted an experiment on a Costas phase-locked loop based on the BPSK signal. When the phase difference between the in-phase branch signal and the quadrature branch signal output by the balance detector was 90°, the signal light and local oscillator light could be locked quickly. This experiment could better demodulate the baseband signal in the input signal light.

In 2020, Wang [95] designed a multi-stage compound-loop control system based on a Costas optical phase-locked loop and completed a BPSK signal-demodulation

experiment with a communication rate of 10 Gbit/s. The system had a bit-error rate better than 10^{-7} and the system had fast-acquisition capability.

In 2020, He [96] established a closed-loop acquisition model of a homodyne coherent-optical communication system, and experimentally measured the frequency-discrimination range as ±2.5 GHz at a communication rate of 5 Gbit/s.

In 2021, Wang [97] designed a balanced optical phase-locked loop for distributed-feedback semiconductor lasers, based on real-time data acquisition and identification. The experimental results showed that loop locking reduced the laser linewidth from 4 MHz to 20 kHz.

In 2021, Huo [98] adopted homodyne laser phase-locked loop technology to conduct coherent-detection experiments on laser-pulse echoes. The results showed that the loop reduced the initial frequency difference between the pulse laser and local oscillator laser to less than 2 MHz for a differential frequency signal within the range of 110 MHz. In addition, the loop-capture range was 112 MHz and the loop bandwidth was 539 MHz.

In 2022, Zhang et al. [99] studied a high-stability laser-frequency stabilization method, based on an optical phase-locked loop. They used the optical phase-locked loop system to investigate the frequency-offset locking of a tunable external-cavity semiconductor laser to the optical frequency comb. After the loop locking, the beat frequency fluctuated within the range of ±0.3 Hz; after the system locking, the beat-frequency linewidth was compressed from 500 kHz to 2 kHz.

In 2022, Fang et al. [100] proposed to use optical phase-locked loop technology in coded synthetic-aperture laser radar to suppress the random phase noise of the system itself. The system worked at 1550 nm and its modulation bandwidth was 1.25 GHz.

Electronic Phase-Locked Loops

Basic Structure of an Electronic Phase-Locked Loop

A phase-locked loop (PLL) synchronizes the output signal and input reference signal in terms of frequency and phase [16]. When the phase-locked loop is locked, the phase difference between the oscillator output signal and reference signal is constant or zero [101].

Figure 26.1 shows the basic circuit of a phase-locked loop, including three basic units: a phase detector (PD), loop filter (LPF), and voltage-controlled oscillator (VCO). To double the clock frequency, the PLL adds a frequency divider between the output circuit of the voltage-controlled oscillator and the input circuit of the phase detector. It is sometimes added after the reference-frequency clock circuit [102] to obtain more output-frequency points.

The working principle of the electric phase-locked loop is as follows. The phase detector compares the phase of the feedback-clock signal with the reference-clock signal, thus obtaining the phase-deviation information. After the signal is processed by the loop filter, the high-frequency component of the signal is eliminated and a

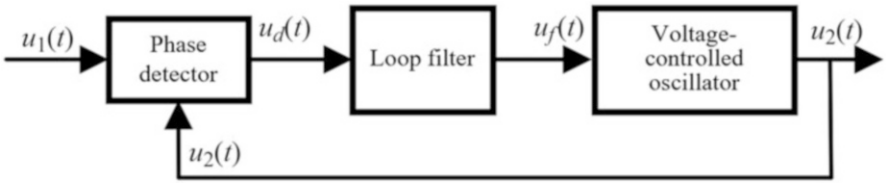


Fig. 26.1 Basic block diagram of a phase-locked loop

voltage signal is generated, which is used to control the frequency of the voltage-controlled oscillator.

One part of the output frequency of the voltage-controlled oscillator is used as the output signal. The other part is used as the feedback-clock signal and is then sent to the phase detector, thus forming a complete phase-locked loop. After the loop stabilizes, the frequency phases of the reference signal and oscillator output signal match [103].

1. Phase detector: The phase detector compares the phase difference between the input signal and the reference signal, and reflects the difference through voltage or current. The average voltage characteristic curve of the output signal is proportional to the phase difference of the input signal.
2. Loop filter: The loop filter retains the DC control signal output by the phase detector of the previous stage and simultaneously removes high-frequency burrs and noise. The number of zeros and poles in the loop affects the loop bandwidth and the phase margin of the phase-locked loop structure.
3. Voltage-controlled oscillator: This module converts the DC-voltage control signal into an AC oscillation signal [104]. This circuit is responsible for generating the output signal. It has an important impact on the output-frequency range of the phase-locked loop and the out-of-band noise.

Classification of Electric Phase-Locked Loops

Different Circuit Structures

Analog Phase-Locked Loops

All parts of an analog phase-locked loop are analog circuits. The analog phase-locked loop is also called a linear phase-locked loop (LPLL) because in the ideal state, when the phase-locked loop works stably, each component-circuit unit works linearly [105]. The loop characteristics of analog phase-locked loops are mainly determined by the characteristics of the phase detector, which plays an important role in signal modulation and demodulation.

Figure 26.2 shows the basic structure of an analog phase-locked loop, which mainly consists of five analog-circuit modules: a phase-frequency detector (PFD),

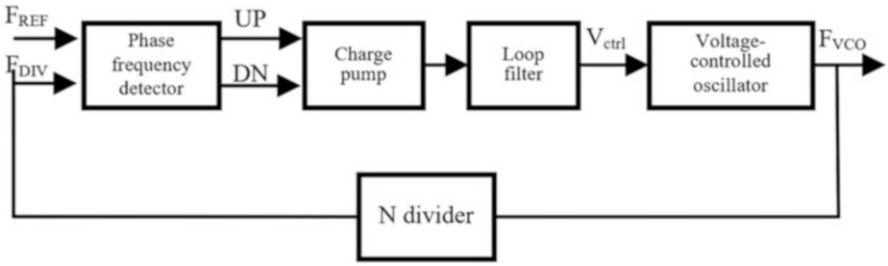
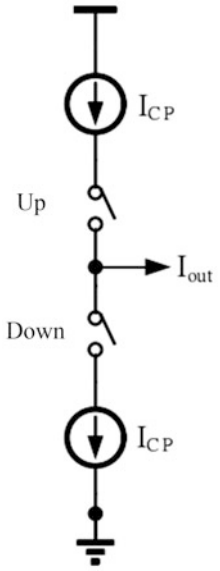


Fig. 26.2 Analog phase-locked loop structure

Fig. 26.3 Schematic diagram of the charge pump



charge pump (CP), loop filter (LPF), voltage-controlled oscillator (VCO), and frequency divider.

In an analog phase-locked loop, the frequency divider divides the output signal. The charge pump converts the output signal of the previous stage into an electric charge and charges and discharges the capacitor in the next stage. Figure 26.3 shows a schematic diagram of the charge pump. The circuit uses two switches to control the current source. The switch signal is the output signal of the frequency and phase detectors. The charge pump directly affects the performance of the entire phase-locked loop.

Following is the working process of the analog phase-locked loop shown in Fig. 26.2. The phase detector compares the phases of the input signal and the output

signal of the frequency divider to generate UP and DN signals, which adjust the charge pump to charge or discharge the loop filter. The high-frequency noise in the output signal of the charge pump is removed in the loop filter. Then, a stabilized signal is generated to adjust the frequency of the output signal of the voltage-controlled oscillator to lock the phases. This is repeated until the entire loop is completely locked; that is, the phase-locked loop system as a whole can be regarded as a negative-feedback system.

All-Digital Phase-Locked Loops

An all-digital phase-locked loop does not contain passive components. Its structure is shown in Fig. 26.4. In the all-digital phase-locked loop, the converter replaces the frequency and phase detector and the charge-pump unit in the traditional phase-locked loop. The converter compares the phases of the reference clock and the feedback clock, obtains the phase-difference signal, and converts it into a digital signal.

The converter in the all-digital phase-locked loop has three different structures: a phase-to-digital converter (PDC), frequency-to-digital converter (FDC), and time-to-digital converter (TDC). The phase-error quantization function of a time-to-digital converter is unique to all digital phase-locked loops. The time-to-digital converter can effectively convert the time to a digital signal, and the clock information can be easily extracted [106]. However, it is difficult to extract from the phase-to-digital and frequency-to-digital converters.

A digital loop filter (DLF) replaces the loop filter. Its function is the same as that of the loop filter in a traditional phase-locked loop. It processes the digital signal output by the converter, filters the noise and high-frequency components, generates the control word, and sends it to the next oscillator. The digital filter directly affects the locking speed and stability of the entire system; therefore, the performance of the loop can be changed by adjusting the parameters of the filter.

The all-digital phase-locked loop oscillator differs from the traditional phase-locked loop oscillator. Its control signal is a digital signal rather than an analog voltage; therefore, it is called a digitally controlled oscillator (DCO). The DCO receives the phase-difference signal generated from the time-to-digital converter to change the output oscillation frequency. The DCO generates an output signal with a specific frequency range and precisely locks its frequency through a phase-locked loop.

Classification of All-Digital Phase-Locked Loops

At present, there are many ways to implement all-digital phase-locked loops; however, the design is primarily based on two types of phase-locked loops [108]. One places the time-to-digital converter in the forward path, which is called a Type-I all-digital phase-locked loop. The other places the time-to-digital converter in the feedback path, which is called a Type-II all-digital phase-locked loop.

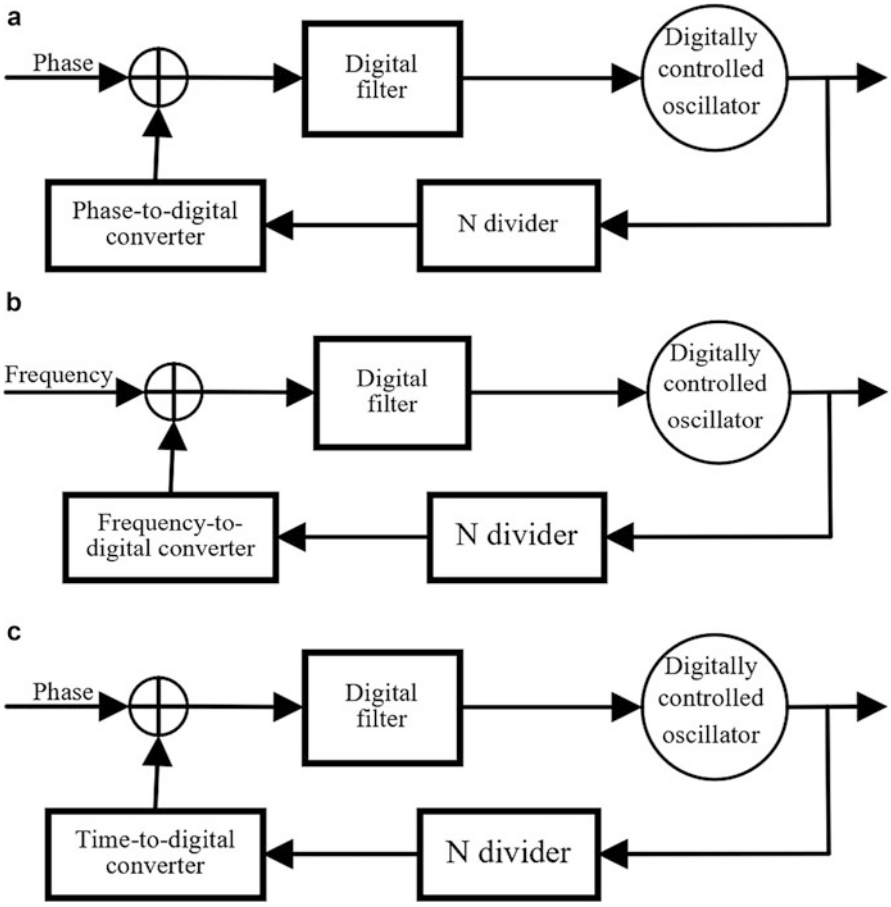


Fig. 26.4 Three all-digital phase-locked loop structures [107]. (**a**) PDC structure (**b**) FDC structure (**c**) TDC structure

Type-I All-Digital Phase-Locked Loop

As shown in Fig. 26.5, the circuit consists of a time-to-digital converter, digital filter, digitally controlled oscillator, and *N* divider. It was developed from the traditional analog phase-locked loop, with mature technology, good dynamic characteristics, and a simple structure. It is applicable to situations in which the frequency or phase must be quickly captured. When the loop is locked, there is a constant phase difference between the reference and output signals. However, the in-band phase noise is high, and the performance is poor.

Following is the working principle of Fig. 26.5. The time-to-digital converter compares the difference between the rising and falling edges of two input signals and quantizes them into the next level of the digital signal input. The digital filter adopts an infinite impulse response (IIR) filter, which is composed of an adder, D trigger, and shift register. Its function is similar to that of the loop filter. After removing the

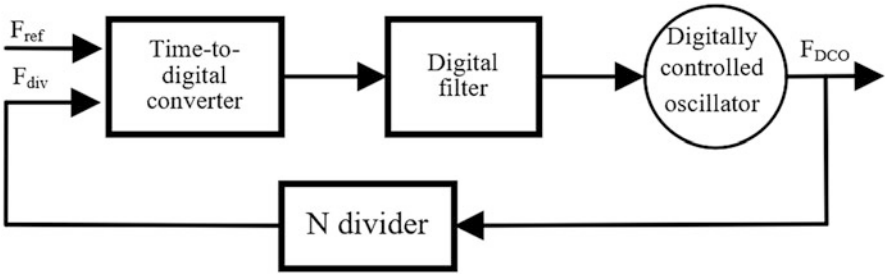


Fig. 26.5 Type-I all-digital phase-locked loop (TDC in forward path) [108]

high-frequency noise, the output signal of the digital filter is used as the input control signal of the digitally controlled oscillator.

The digitally controlled oscillator uses a MOS capacitor to determine the frequency resolution, whereas a voltage-controlled oscillator uses the C-V characteristic of the varactor to realize high-precision tuning. A digitally controlled oscillator uses delta–sigma modulation (DSM), which can reach or even exceed the tuning accuracy of a voltage-controlled oscillator.

The structure of the *N*-frequency divider is generally similar to that of the analog phase-locked loop. It can also be applied to digital circuits. The divided signal is fed back to the time-to-digital converter for comparison, and the entire process is repeated until the all-digital phase-locked loop is completely locked.

Type-II All-Digital Phase-Locked Loop

As shown in Fig. 26.6, the Type-II all-digital phase-locked loop differs from the traditional phase-locked loop in structure. The circuit structure includes a time-to-digital converter, digital integrator, digital filter, digitally controlled oscillator, and high-speed counter. Compared with the Type-I all-digital phase-locked loop, the time-to-digital converter is placed in the feedback path to measure the phase difference between the rising or falling edge of the digitally controlled oscillator and the reference clock, which is then converted to digital.

The phase difference between the output signal and the reference signal of the Type-II PLL is approximately zero, and it is often used in high-precision frequency synthesis and clock recovery fields. Its disadvantage is that there is significant quantization noise. Because the input signal of the high-speed counter and time-to-digital converter is the output frequency of the digitally controlled oscillator, the output of the digitally controlled oscillator consumes a large current and high power, and the circuit design and implementation of this type of phase-locked loop are relatively complex.

Following is the working principle of Fig. 26.6. A high-speed counter is included in the feedback channel. When the reference-clock edge arrives, the counter records the integral part of the output-clock phase of the digitally controlled oscillator, and the time-to-digital converter records the fractional part of the clock phase. Thus, the overall digital phase signal of the digitally controlled oscillator is obtained by adding the output of the counter and the time-to-digital converter.

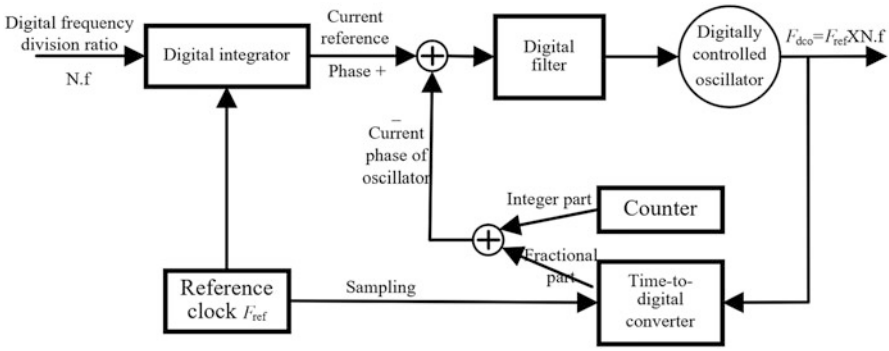


Fig. 26.6 Type-II all-digital phase-locked loop (TDC in feedback channel) [108]

The phase of the reference clock is also converted into a digital phase signal by a digital integrator, and the digital value of the current phase difference is then calculated by comparing the two output digital signals. This value is then passed into the digital filter to filter the noise and is finally used for the output-frequency control of the digitally controlled oscillator.

Digital–Analog Hybrid Phase-Locked Loop

Owing to a further reduction in the process size, the internal-leakage current of the component increases, and the periodic vibration of the clock frequency of the analog phase-locked loop also increases. For leakage current, long-period jitter, and other problems, the all-digital phase-locked loop can be used; however, the resolution of the digitally controlled oscillator in the all-digital phase-locked loop is limited, and there is high flicker noise, which increases jitter; thus, a digital–analog hybrid phase-locked loop is generally used.

To reduce flicker noise in the digital-circuit structure, the fractional frequency division is generally used to replace the digitally controlled oscillator in the digital–analog hybrid phase-locked loop structure [109].

A charge-pump phase-locked loop (CPPLL) is a typical digital–analog hybrid phase-locked loop. Its characteristics include a wide linear range, short acquisition time, and small lock-phase difference. The charge-pump phase-locked loop is shown in Fig. 26.7. Its structure consists of a frequency and phase-frequency detector (PFD), charge pump (CP), loop filter (LPF), and voltage-controlled oscillator (VCO). Compared to the analog phase-locked loop, the only difference is that the phase detector unit uses a digital charge-pump phase detector.

The charge-pump phase-locked loop can filter the jitter from the output of the all-digital phase-locked loop without considering the impact of the leakage current. The jitter of the digital–analog hybrid phase-locked loop is also improved in the clock-output cycle. The disadvantage of the charge-pump phase-locked loop is that the anti-interference performance is poor, and it is vulnerable to noise and other factors. Moreover, its compatibility with the digital module is poor, so its performance level is not too high.

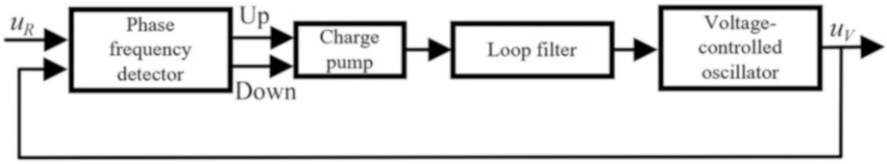


Fig. 26.7 Charge-pump phase-locked loop

Software Phase-Locked Loops

The software phase-locked loop is implemented using software [11]. The parameter design is more flexible and the structural design is simpler. It has better versatility than the other phase-locked loops. The principle of the software phase-locked loop does not differ significantly from that of other phase-locked loops; however, there are differences in the mathematical models and implementation methods. The software phase-locked loop does not have the complex circuit structure of other phase-locked loops but uses microsecond computing time to replace the electronic components in other phase-locked loops. The higher the precision of the software phase-locked loop, the higher the performance, and the more complex the corresponding algorithm instructions.

Its speed and accuracy can be compared with other phase-locked loops, but only if the algorithm instructions can be executed sufficiently fast on the hardware. If the algorithm instructions run slowly on the hardware platform or the hardware platform is not sufficiently powerful, then a more powerful hardware platform needs to be adopted, which will also increase the cost. Therefore, the software phase-locked loop is applicable to systems with low and low-frequency requirements.

Comparison of Analog and Digital Phase-Locked Loops

Phase-locked loops are classified into analog and digital. Table 26.2 compares their performance.

The analog phase-locked loop has a strong tracking ability for signals with sinusoidal characteristics, and its locking time is very short, ensuring a steady-state difference between the output clock and reference clock. The voltage-controlled oscillator limits the center frequency of the analog phase-locked loop to a small range. The performance of the analog phase-locked loop depends on the transistor size. The expected performance is achieved using a complex circuit structure and sacrificing power consumption. The loop-filter capacitance occupies a large area. With the continuous reduction in the process size, the operating-frequency range of the analog phase-locked loop will be significantly reduced.

The digital phase-locked loop has strong scalability, good portability, excellent anti-interference performance, and is easy to integrate in large-scale digital circuits [110]. The actual circuit often works in an imperfect environment, and the analog phase-locked loop circuit is easily affected by external imperfect factors, whereas the digital phase-locked loop circuit effectively overcomes the problems caused by changes in the process, voltage, temperature, etc. [111].

Table 26.2 Comparison between analog and digital phase-locked loops

Performance metric	Analog phase-locked loop	Digital phase-locked loop
Lock time	Longer	Shorter
Lock speed	Faster	Slower
Noise resistance	Weak	Stronger
Circuit complexity	Higher	Lower
Portability	Poor	Better
Area	Larger	Smaller
Power dissipation	Same	Same
Jitter performance	Better	Related to control word
Compatibility	Digital circuit needs to be considered	No need to consider in digital circuit

Other Classifications

Other classifications are based on analog and digital phase-locked loops.

Phase-Locked Loops with Different Frequency-Division Coefficients

According to the different frequency-division coefficients of the frequency divider, phase-locked loops can be divided into integer frequency-division and fractional frequency-division.

When the phase-locked loop is unlocked, the frequency and phase detector compares the reference signal F_{ref} and the divided signal F_{out} and outputs the phase-error signal between them. After feedback, the loop can lock the reference and feedback signal phases. Therefore, the reference-signal frequency must be equal to the frequency of the signal output after frequency division; thus, there is a locking relationship between the output frequency of the voltage-controlled oscillator and N times the input reference frequency, as shown in Eq. (26.1):

$$F_{\text{out}} = N \times F_{\text{ref}}. \tag{26.1}$$

Figure 26.8 shows an integer frequency-division phase-locked loop. The entire loop has a simple structure and good noise performance. The disadvantage of this method is that the frequency resolution of the integer phase-locked loop is low. When the circuit has high accuracy requirements, it cannot meet the design requirements.

Because N is an integer, the minimum step length of the phase-locked loop is equal to the frequency F_{ref} of the reference signal. Therefore, to obtain a lower step length, the input reference frequency must be reduced; however, this will increase the value of the frequency divider and increase the in-band noise accordingly. Consequently, the overall performance of the phase-locked loop is reduced.

To improve the in-band noise of the integer frequency-division phase-locked loop, the value of the frequency divider can be reduced. However, when the output frequency is the same, the input reference frequency must increase; that is, the frequency and phase detector must have a higher phase-discrimination capability.

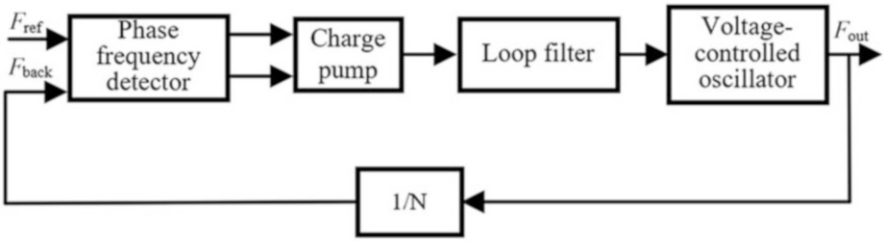


Fig. 26.8 Structure block diagram of an integer frequency-division phase-locked loop [112]

Therefore, when the output frequency requires a higher accuracy, the fractional frequency-division phase-locked loop is generally selected. As shown in Fig. 26.9, the fractional-frequency phase-locked loop [112] is generally constructed based on the charge-pump phase-locked loop structure.

The PLL in Fig. 26.9 consists of a frequency and phase-frequency detector (PFD), charge pump (CP), loop filter (LPF), voltage-controlled oscillator (VCO), and a two-mode frequency divider controlled by a delta–sigma modulator. The delta–sigma modulator consists of a delta modulator and an integrator that acts as a summation. With this structure, it is possible to obtain a higher frequency resolution and push low-frequency noise into high-frequency noise to shape or modulate the noise [113].

In a fractional-frequency phase-locked loop, the delta–sigma modulator quantization error can be removed by a low-pass filter, and the influence of the quantization error on the output phase noise can be reduced to a negligible degree by reducing the bandwidth of the phase-locked loop.

Phase-Locked Loops of Different Types and Numbers

Phase-locked loops are divided into different types according to the number of integrators in the loop, that is, according to the number of zeros. Many important properties of phase-locked loops are related to their type number. Because the voltage-controlled oscillator performs an integral operation, the phase-locked loop is of at least one type. In contrast, Class-II phase-locked loops are widely used, whereas Class-III phase-locked loops are occasionally used, owing to their complex structure.

Narrowband and Broadband Phase-Locked Loops

Bandwidth is the most basic property of a phase-locked loop. Depending on the bandwidth, they can be divided into narrowband and broadband. However, their applications are different.

The narrow band and automatic-tracking characteristics of narrowband phase-locked loops make them widely used in receivers. In the receiver, the frequency or phase contains both the input-signal information and additive noise. The phase-locked receiver correctly reconstructs the original signal to eliminate as much additive noise as possible. When the narrowband phase-locked loop inputs a signal with a stable frequency, the local oscillator can track it with less information. The phase-locked loop in this application can be regarded as a device for filtering noise

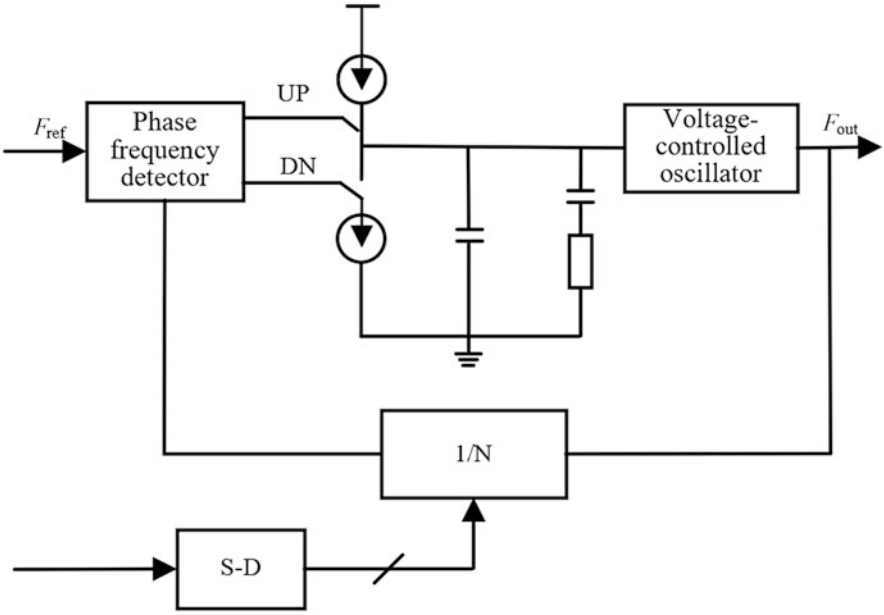


Fig. 26.9 Traditional fractional frequency-division phase-locked loop [112]

[114]. The narrowband phase-locked loop is primarily used for carrier synchronization and reducing the clock jitter of digital circuits. The input signal can be tracked when the bandwidth of a broadband phase-locked loop is sufficiently wide.

In a frequency-modulation system, the voltage of the frequency-modulation signal directly controls the modulation signal output by the voltage-controlled oscillator. In practice, to obtain high-quality modulated signals, open-loop phase-locked loop frequency modulators, closed-loop phase-locked loop frequency modulators, and two-point modulators are typically used to effectively suppress the impact of the nonlinearity of the voltage-controlled oscillator.

Figure 26.10 shows the structure diagram of a closed-loop phase-locked loop frequency modulator. In this structure, the transfer function from the loop-filter noise to the output-phase noise is consistent with the transfer function from the modulated signal to the output. Figure 26.11 shows the structure diagram of phase-locked loop signal demodulation, where the control voltage is the demodulated output signal.

Parameters of Electronic Phase-Locked Loops

Following are the parameters of electronic phase-locked loops.

1. Frequency tuning range: This refers to the range between the minimum and maximum output frequencies of the phase-locked loop, that is, the size of the

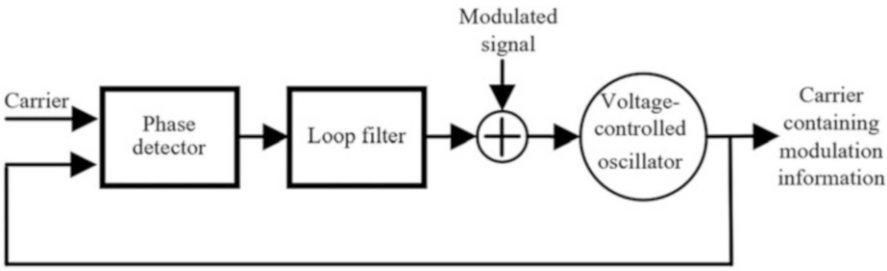


Fig. 26.10 Closed-loop phase-locked loop frequency modulator [115]

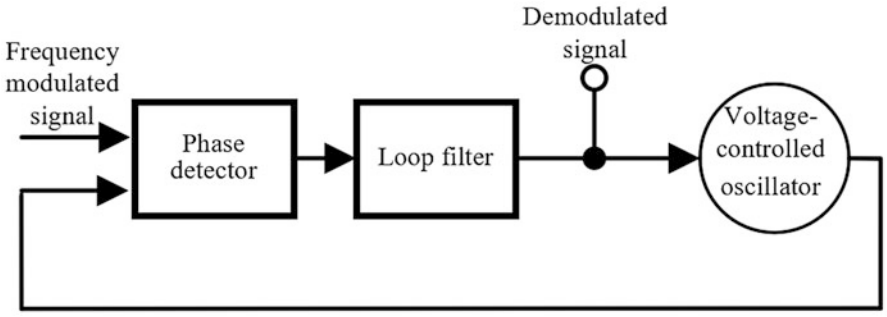


Fig. 26.11 Signal demodulation using a phase-locked loop

output frequency. Generally, the best tuning range of the output frequency and the appropriate frequency band are selected according to the communication standard and specific situation.

2. Frequency resolution: This refers to the minimum frequency interval of the signal output by the phase-locked loop.
3. Frequency conversion time: The time required for the phase-locked loop to change from any frequency to another frequency must meet this requirement for the communication system to work in different frequency bands.
4. Phase noise and spurious components: If phase noise is taken as the measurement of the output frequency, it is ideal for the PLL to output a single frequency point pulse in the frequency domain. However, in practical applications, owing to noise and various non-ideal effects, the output signal contains other frequency components that are generally called spurious.
5. Lock time: When the phase-locked loop is in a stable operation state, it is said to have entered the lock state. The requirements for the length of the lock time differ in different wireless communication systems. The loop bandwidth significantly affects the phase lock time [74]. When the loop bandwidth is less than one-tenth of the reference-clock signal frequency, the loop stability is good.

Applications of Electric Phase-Locked Loops

In a pulse-position modulation (PPM) and demodulation optical wireless communication system, the receiver-information detection part of the optical wireless communication system must have a strict time-synchronization relationship with the superframe to ensure that the communication system can normally demodulate the information in the PPM frame.

Figure 26.12 shows the structure of an optical wireless communication receiver. At the receiver end, after the monopolar narrow pulse signal is output by the avalanche photodetector (APD), it reaches the preprocessing part, and the high-frequency noise in the signal is removed through the low-pass filter. Finally, the received signal is sent by the splitter to the PPM detection demodulator, slot-synchronization unit, and frame-synchronization unit for corresponding processing.

The time-slot synchronization unit of an optical wireless communication receiver is shown in Fig. 26.13. It primarily uses a digital phase-locked loop to extract the signal of the time-slot clock. First, the interference noise with lower amplitudes of the input narrow pulse signal is filtered in the amplitude comparator. Then, the signal is processed by the width comparator for noise reduction. The pulse signal is transmitted to the pulse stretcher, so that the duty cycle of the signal is adjusted to 50%. The corresponding level signal is output, then sampled by the analog-to-digital converter (A/D), and finally sent to the digital phase-locked loop. The clock signal is

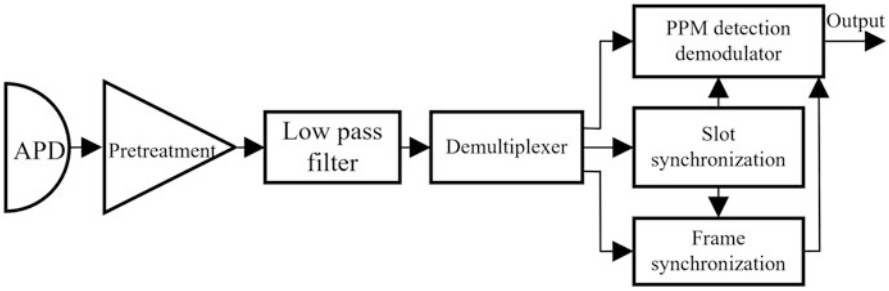


Fig. 26.12 Block diagram of an optical wireless communication receiver [116, 117]

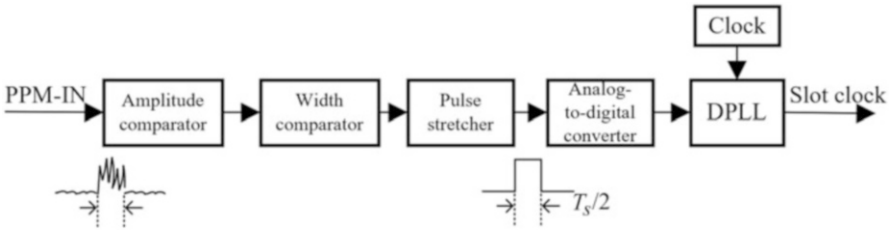


Fig. 26.13 Time-slot synchronizer [116, 117]

used to extract the synchronous-slot clock and ensure slot synchronization in optical wireless communication.

Electric phase-locked loop technology also plays an important role in radio and television, measuring instruments and meters, remote control, national defense technology, electronic countermeasures, digital signal processing, mechanical processing, and other fields.

Optical Phase-Locked Loops in Optical Communication

Basic Structure of an Optical Phase-Locked Loop

The optical phase-locked loop is essentially the same as the traditional electric phase-locked loop. It is also a closed-loop locking circuit for the frequency and phase. In particular, an optical phase-locked loop is a control system that acts in the optical frequency range.

The structural difference between the two is that the phase-detection function of the optical phase-locked loop is completed through the optical hybrid and the front-end circuit. It uses a tunable laser to replace the voltage-controlled oscillator in the electric phase-locked loop, and adds a photoelectric detection circuit after the optical hybrid. It converts the differential frequency signal output by the optical hybrid into an electrical signal so that the loop-filter circuit can perform the corresponding processing [96]. Therefore, in the theoretical analysis, the optical and electric phase-locked loops are the same.

Figure 26.14 shows the basic circuit structure of the optical phase-locked loop. It is composed of four units: an optical hybrid, photoelectric detection circuit, loop filter, and tunable laser.

Following is the working principle of the structure shown in Fig. 26.14. First, in the optical hybrid, the input signal light and the local oscillator light are mixed coherently. Then, the mixed light is converted into the corresponding electrical signal through the photoelectric detection circuit, filtered, and integrated into the loop filter. This generates a signal in the tunable laser, which is used to control the optical frequency output by the local oscillator laser. Then, the output signal of the local oscillator laser is fed back to the optical hybrid. This is repeated to achieve a closed-loop control. During this process, the output frequency of the local oscillator laser is constantly close to the input-signal frequency.

Optical Hybrids

The function of the optical hybrid is to couple the input signal light and local oscillator light, according to a certain phase relationship to produce the mixing light. Optical hybrids are mainly divided into 90° optical hybrids and 180° optical hybrids. Their equivalent phase model is shown in Fig. 26.15.

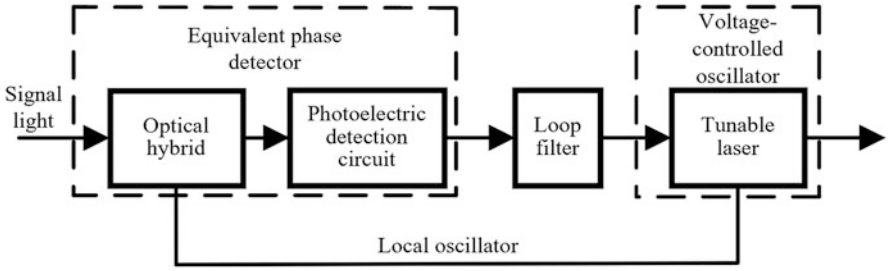


Fig. 26.14 Basic structure of an optical phase-locked loop

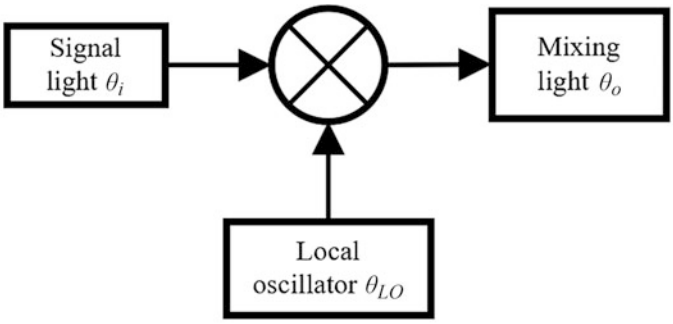


Fig. 26.15 Equivalent phase model of an optical hybrid (θ represents the phase)

In Fig. 26.15, setting the signal light as $u_i = U_i \sin(\omega_i t + \theta_i)$ and the local oscillator as $u_{LO} = U_{LO} \sin(\omega_{LO} t + \theta_{LO})$, the mixing light can be expressed as

$$u_o = K_d \sin(\theta_o), \tag{26.2}$$

where $\theta_o = \theta_i - \theta_{LO}$, $K_d = \frac{1}{2} K_m U_i U_{LO}$, and K_d is the maximum output of the optical hybrid's equivalent phase detector.

Photoelectric Detection Circuit

The photoelectric detector converts the optical phase-error signal output by the optical hybrid to the corresponding electrical signal. Generally, a balanced detector is used because it can better suppress the laser noise. Figure 26.16 shows the structure of the balanced detector. A balanced detector means that the two photoelectric detectors have the same performance, as far as possible. After subtraction, the signal is doubled, the noise is eliminated, the direct-current component disappears, and the signal-to-noise ratio is greatly improved.

The photoelectric-detection circuit is fabricated using the photoelectric effect of certain substances in the semiconductor material under light radiation. Electrons in the semiconductor material are excited by photons to form a current. The other type of detector uses matter to induce a photothermal effect under light irradiation. The photon energy acts on the lattice and causes the temperature to rise and enhances the vibration, which changes the characteristics of matter and forms a current.

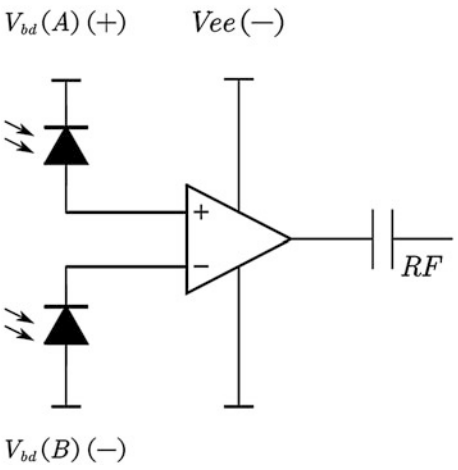


Fig. 26.16 Balanced detector structure

Loop Filter

While filtering the high-frequency components and noise in the signal, the loop filter can also affect the loop stability, acquisition time, loop-acquisition bandwidth, and tracking time by changing its parameters.

The working principle is to use an integral circuit to convert the photodetector output into the corresponding error-feedback voltage. While effectively adjusting the loop gain and compensating the loop-phase margin, the low-pass filter circuit further reduces the loop noise.

The key to designing an optical phase-locked loop is the open-loop transfer function. The transfer function of the phase-detector unit and the tunable laser in the optical phase-locked loop is determined by the characteristics of the device itself, and the design space is small. Therefore, the loop filter is the key to designing an optical phase-locked loop, and is equivalent to the control algorithm in a general control system.

In the Laplace-transform domain of the basic phase-locked loop, $\theta_i(s)$ is the input signal. $\theta_o(s)$ is the output signal of the voltage-controlled oscillator. When the loop is in the locked state, the equivalent phase-detector output voltage of the optical hybrid is shown in Fig. 26.17:

$$V_d = K_d(\theta_i - \theta_o). \tag{26.3}$$

We set the phase error $\theta_e(s)$ as

$$\theta_e = \theta_i - \theta_o. \tag{26.4}$$

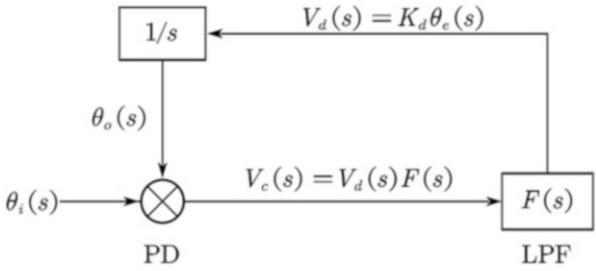


Fig. 26.17 Basic block diagram of a phase-locked loop

We assume that V_c is the control voltage used to control the frequency of the voltage-controlled oscillator, expressed as

$$V_c(s) = V_d(s)F(s). \tag{26.5}$$

The input control voltage of the voltage-controlled oscillator is $V_c(t)$. The frequency offset is K_oV_c. Subsequently, the output signal of the voltage-controlled oscillator is transformed into

$$\theta_o(s) = \frac{K_oV_c(s)}{s}. \tag{26.6}$$

The open-loop transfer function of the phase-locked loop is

$$G(s) = \frac{\theta_o(s)}{\theta_e(s)} = \frac{K_dK_oF(s)}{s}. \tag{26.7}$$

The closed-loop transfer function is

$$H(s) = \frac{\theta_o(s)}{\theta_i(s)} = \frac{G(s)}{1+G(s)} = \frac{K_dK_oF(s)}{s+K_dK_oF(s)}. \tag{26.8}$$

The error transfer function is

$$E(s) = \frac{\theta_e(s)}{\theta_i(s)} = \frac{1}{1+G(s)} = 1 - H(s) = \frac{s}{s+K_dK_oF(s)}, \tag{26.9}$$

where $s = j\omega$. $F(s)$ is the transfer function of the loop filter; it is determined by the specific structure of the filter. From the above expression, we can see that the system transfer function performs low-pass filtering on the input signal and the error transfer function performs high-pass filtering on the input signal.

Tunable Laser

The linewidth, noise intensity, tunable range, and other parameters of the tunable laser directly affect the optical phase-locked loop, which can be regarded as a linear voltage-frequency conversion device. The instantaneous angular frequency is expressed as follows:

$$\omega_{LO}(t) = \omega_o + K_l V_c(t), \tag{26.10}$$

where ω_o is the initial angular frequency, $V_c(t)$ is the input control voltage of the tunable laser, and K_l is the gain coefficient of the tunable laser ($rad \cdot Hz/V$).

Equation (26.10) is expressed as a Laplace transfer operator in the form of

$$K_{LO}(s) = \frac{k_l}{s}. \tag{26.11}$$

Classification of Optical Phase-Locked Loops

Initially, the optical phase-locked loop was developed based on the electric phase-locked loop. With the emergence of coherent communication technology, optical phase-locked technology has developed rapidly and various optical phase-locked loops have emerged. The following is a comparative analysis of the composition principle and working performance of optical phase-locked loops.

Balanced Optical Phase-Locked Loop

Figure 26.18 shows the balanced phase-locked loop. Its structure is simple and easy to implement, and is the basis of other optical phase-locked loops. It is composed of a 180° optical hybrid, a balance detector, a loop filter, and a local oscillator laser. It has a symmetrical structure. The local oscillator and signal light are coherently mixed by a 180° optical hybrid to generate two optical signals with a phase difference of 180°. The optical signal is converted into an electrical signal using a photoelectric balance detector.

The direct-current component of the electrical signal is filtered by a trans-resistance amplifier, and the electrical signal containing the phase error of the two light beams is filtered and corrected by a loop filter to generate a signal to control the local oscillator laser. When the system is in a closed-loop state, the frequencies of the local oscillator light and signal light will be close to each other, and the phases of the local oscillator light and signal light are synchronized.

The balanced phase-locked loop must have residual carriers perpendicular to the local oscillator light so that it can use the current signal generated by the phase shift to control the output of the local oscillator laser and lock the phase. In addition, owing to the presence of residual carriers, the loop can output baseband signals only after secondary demodulation. The balanced phase-locked loop can transmit analog signals and process digital signals. Using a 180° optical hybrid can minimize the loss from power consumption. The laser linewidth is related to the communication rate, whereas the relationship between the maximum linewidth of the balanced phase-locked loop laser and the system-transmission rate is $\Delta\nu \leq 5.8 \times 10^{-6} R_b$ [118], where R_b is the bit rate in bit/s, when the bit-error rate is 10^{-10} and the power loss is 1 dB.

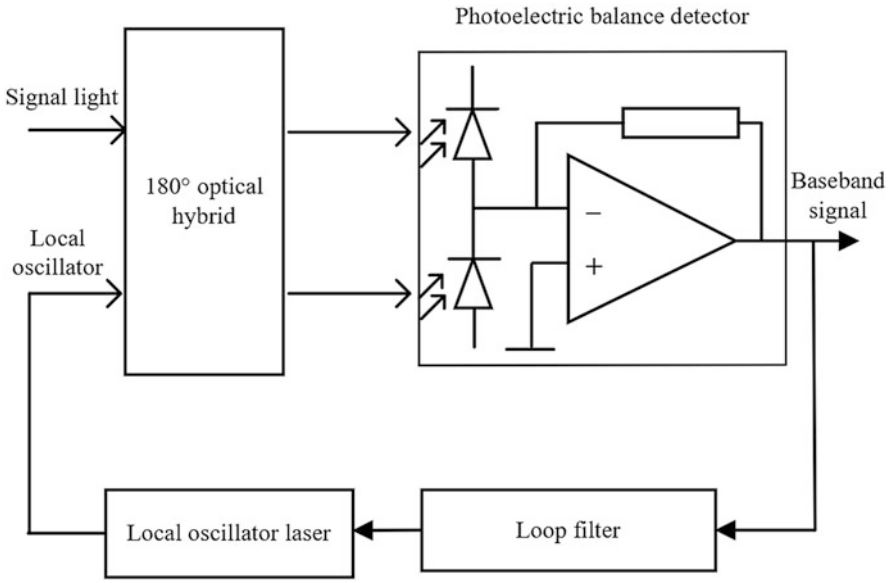


Fig. 26.18 Balanced optical phase-locked loop [118]

Costas Optical Phase-Locked Loop

The Costas loop uses a 90° optical hybrid, whose structure is shown in Fig. 26.26. In the 90° optical hybrid, the local light and signal light are coherently mixed, and then four optical signals with a phase difference of 90° are generated. After conversion by the photoelectric balanced detection circuit, the four optical signals are divided into two signal channels, namely, the in-phase and quadrature channels. These two channels carry modulation information with a phase difference of 90°. The signals from these two channels are multiplied to eliminate modulated signals.

In this process, a phase-error signal of the local oscillator and signal light with twice the frequency are generated, and the signal is processed by the loop filter to obtain a voltage signal, which controls the output of the local oscillator laser. The frequencies of the local oscillator light and the signal light continue to approach each other, and the intermediate-frequency signal continues to decrease until the frequency phases of the local oscillator light and the signal light are equal; that is, the voltage in the loop is zero, so that the loop is locked. At this time, the output of the same branch is the baseband signal (Fig. 26.19).

Because the Costas loop uses a 90° optical hybrid with a complex structure, it does not need to transmit residual carriers, nor does it have the problem of crosstalk between the phase-locked signal and data. An alternating-current coupling element is used in the loop, which effectively suppresses the noise interference generated by the laser; therefore, the Costas loop has high reception sensitivity.

Compared to the balanced phase-locked loop, the structure of the 90° optical hybrid is much more complex than that of the 180° optical hybrid, and its

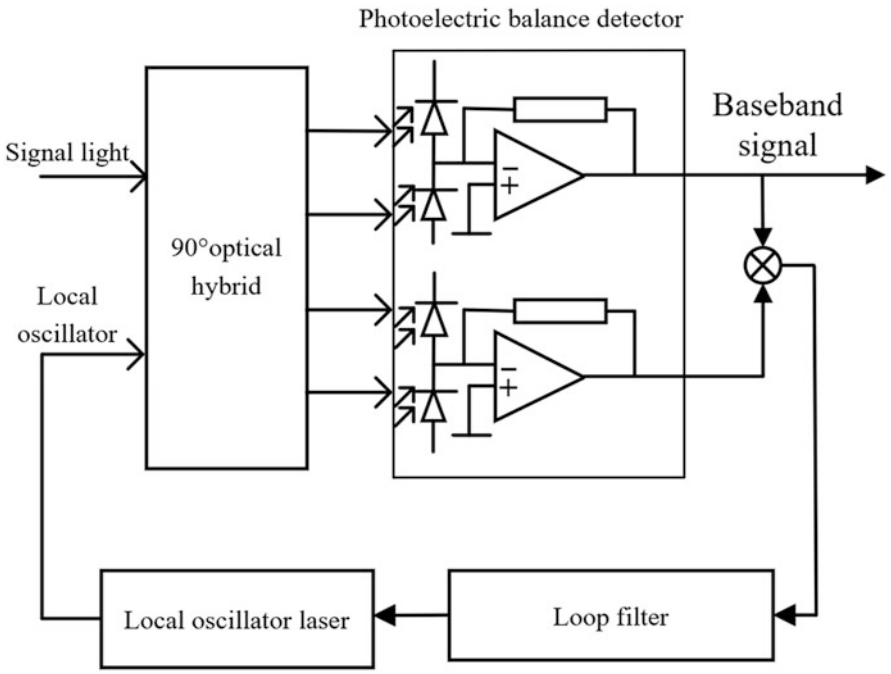


Fig. 26.19 Costas optical phase-locked loop [65]

implementation is more difficult. The Costas loop can process both analog and digital signals, and its laser-linewidth requirements are relatively low. The relationship between the maximum linewidth of the laser and the system transmission rate is $\Delta v \leq 5.02 \times 10^{-4} R_b$ [65], where R_b is the bit rate in bit/s, when the bit-error rate is 10^{-10} and the power loss is 1 dB.

Decision-Driven Optical Phase-Locked Loop

A decision-driven optical phase-locked loop is an optical phase-locked loop that suppresses the carrier wave. The structure is shown in Fig. 26.23. The circuit includes a 90° optical hybrid, a photodetector, a loop filter, a delayer, a threshold comparator, and a local oscillator (LO) laser. Comparing the decision-driven optical phase-locked loop with the Costas optical phase-locked loop, it can be found that the two have considerable structural similarity. Only the data-decision circuit unit is added to the in-phase channel of the decision-driven phase-locked loop; it is composed of a low-pass filter circuit and a threshold comparator.

A one-bit delay circuit is added to the quadrature channel, which can compensate for the problem in which the two-channel signals are not synchronized because of the data decision. Although a one-bit delay circuit is added, it is still difficult to ensure that the quadrature branch and in-phase branch are synchronized in time, which affects the signal modulation (Fig. 26.20).

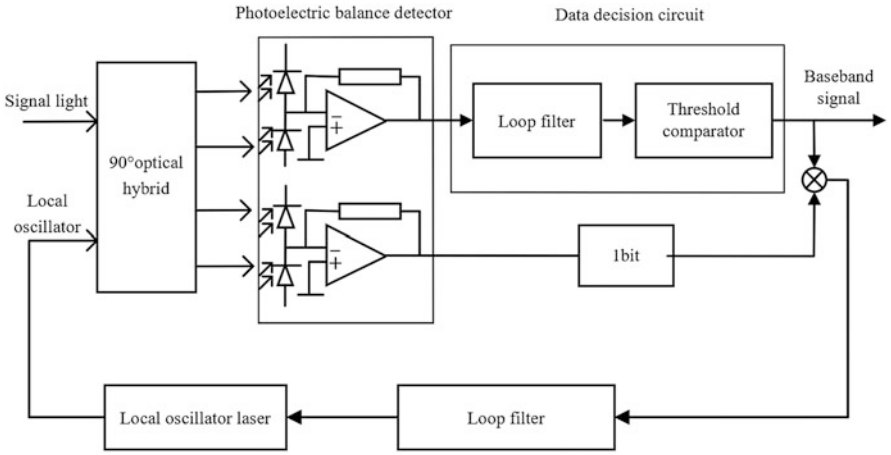


Fig. 26.20 Decision-driven optical phase-locked loop [23]

Compared with the Costas loop, the phase error of the decision-driven optical phase-locked loop is smaller; however, it can only process digital signals. It has higher requirements for the laser linewidth than the Costas loop. The relationship between the maximum linewidth of the laser and the system transmission rate is $\Delta v \leq 4.5 \times 10^{-4} R_b$ [23]. R_b is the bit rate in bit/s, when the bit-error rate is 10^{-10} and the power loss is 1 dB.

SyncBit Optical Phase-Locked Loop

The SyncBit optical phase-locked loop has the advantages of a balanced phase-locked loop and a Costas phase-locked loop. Its structure is simple. The circuit is shown in Fig. 26.24. It involves adding a signal amplifier, data holding, and phase modulation to the balanced phase-locked loop circuit structure. The phase-locked loop uses a 180° optical hybrid to optically mix the signal. It inputs the obtained optical signal to the photoelectric balance detector, and then sends it to the phase-locked circuit after a photoelectric conversion.

In most cases, the phase-locked loop is in an open state and its output signal is used for data demodulation. In this case, the loop can be regarded as an in-phase channel in the Costas phase-locked loop. During data transmissions, the loop is closed and the phase is locked. The output end of the LO laser uses the phase modulator to adjust the phase of the LO light, so that the initial phase of the light wave increases by 90°. This achieves the goal of a 90° phase difference between the optical phase of the output signal of the photoelectric balance detector and the phase of the previous moment. At this time, the loop can be regarded as an orthogonal channel in Costas-loop phase locking.

The phase-error information is obtained by multiplying the electrical signals of the two channels. Then, the loop completes the phase-locking, uses the signal

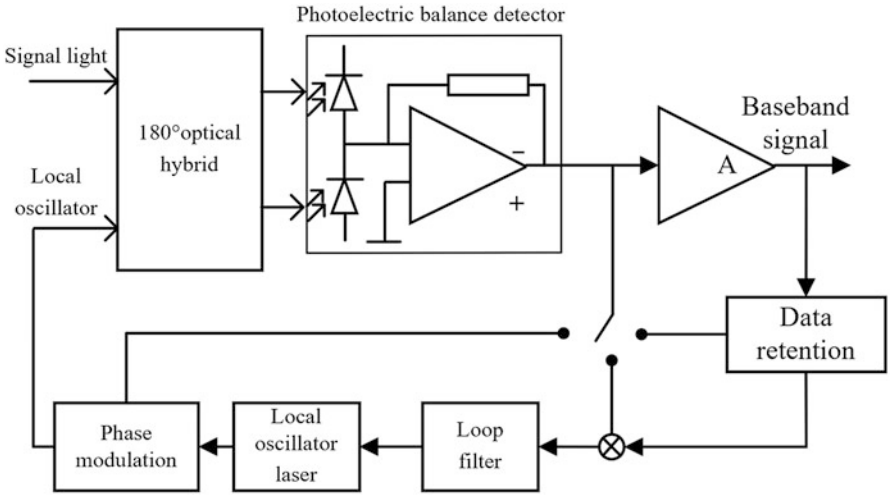


Fig. 26.21 SyncBit optical phase-locked loop [119]

amplifier to amplify and filter, and restores the output of the photoelectric balance detector by the baseband signal (Fig. 26.21).

The advantage of the SyncBit optical phase-locked loop is that the signal input to the 180° optical hybrid is coherently mixed in a time-sharing manner to achieve the effect of 90° optical mixing. The work of the Costas optical phase-locked loop can be completed using a branch, reducing the complexity of the system.

The disadvantage is that the signal-to-noise ratio is reduced and bandwidth will be lost because the phase-locked loop needs to perform special coding processing on the data, which generates a large amount of redundant information. If the communication bandwidth is increased to solve this problem, the noise will increase. If the same bandwidth is used, the transmission rate will be reduced, and the system's requirements for communication bandwidth will be increased.

The relationship between the maximum linewidth of the laser and the system's transmission code rate is $\Delta\nu \leq 3.1 \times 10^{-4} R_b$ [119]. R_b is the bit rate in bit/s, when the bit-error rate is 10^{-10} and the power loss is 1 dB.

Dither Optical Phase-Locked Loop

The structure of the dither optical phase-locked loop is relatively complex; however, it is applicable to a variety of modulation modes. Its structure is shown in Fig. 26.22, including a 180° optical hybrid, balance detector, radio-frequency amplifier, detector, bandpass filter, loop filter, 90° phase shifter, oscillation source, and local oscillator laser.

The circuit uses a 180° optical hybrid to optically mix the signal light and the local oscillator light, converts the photoelectric signal of the mixed light through the balance detector, and simultaneously outputs the electrical signal. The radio frequency amplifier and the detector in the loop of the dither optical phase-locked

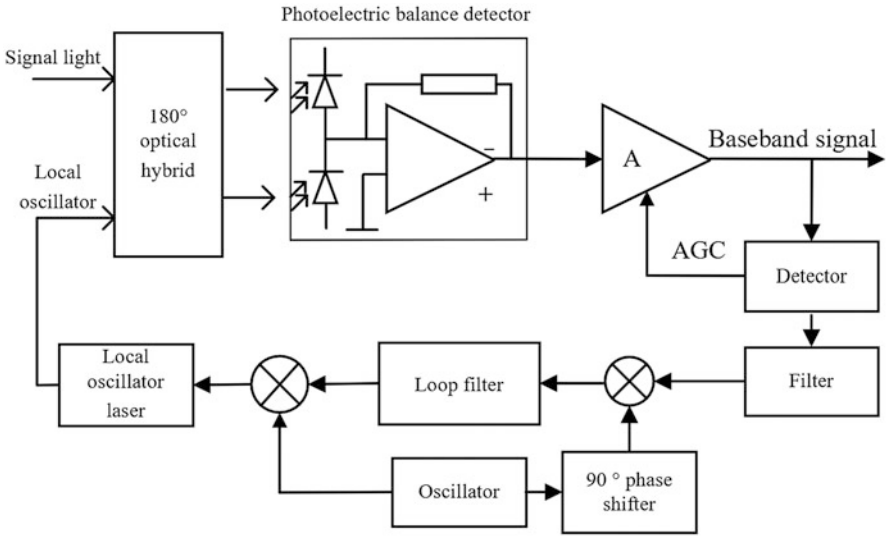


Fig. 26.22 Block diagram of a dither phase-locked loop [43]

loop are combined into an automatic gain control module, which adjusts the electrical signal output from the previous stage and outputs the adjusted electrical signal to the phase lock branch.

The dither optical phase-locked loop is an oscillation source that generates a disturbance in front of the local oscillator laser circuit. After the oscillation source is turned on, the generated disturbance signal is added to the phase of the local oscillator laser. When the loop is unlocked, a frequency signal containing the phase error is detected at the end of the loop, where the signal amplitude represents the magnitude of the phase error. The filtered signal can be used as a control signal to adjust the local oscillator laser and lock the loop.

Compared with the balanced phase-locked loop, the dither optical phase-locked loop can process both analog and digital signals without transmitting the residual carrier. However, additional oscillation signals and corresponding signal processing must be carried out, which reduces the average amplitude of the input signal, resulting in increased loop power consumption. The relationship between the maximum linewidth of the laser and the system transmission rate is $\Delta v \leq 2.6 \times 10^{-4} R_b$ [43]. R_b is the bit rate in bit/s, when the bit-error rate is 10^{-10} and the power loss is 1 dB.

Subcarrier Optical Phase-Locked Loop

The local oscillator laser plays an important role in the negative feedback circuit. Usually, the phase of a local oscillator laser is adjusted by changing the injection current of the laser using a piezoelectric transducer (PZT). However, the response time of the piezoelectric converter limits the laser tuning bandwidth. The key component of the subcarrier optical phase-locked loop (SC-OPLL) circuit structure is the optical voltage-controlled oscillator (OVCO). Figure 26.23 shows a schematic

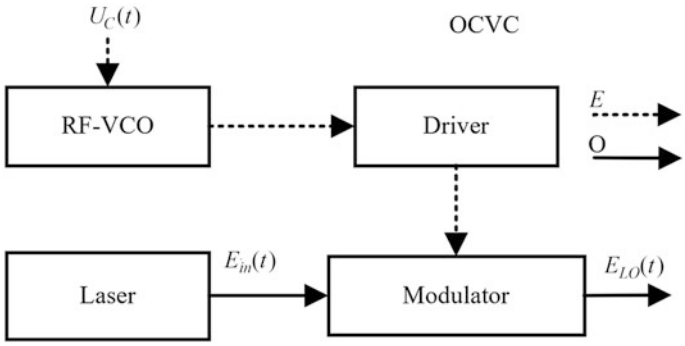


Fig. 26.23 Schematic diagram of an optical voltage-controlled oscillator

diagram. It is composed of a radio-frequency voltage-controlled oscillator (RF-VCO), driver, external modulator, and laser.

The subcarrier optical phase-locked loop does not directly control the tunable laser but modulates the intensity of the local oscillator laser through an optical voltage-controlled oscillator. The nonlinearity of the modulator generates subcarriers. The phase change caused by these subcarriers is caused by a change in their control voltage.

The advantage of the optical voltage-controlled oscillator is that it can change the optical phase of the subcarrier, based on the radio-frequency voltage-controlled oscillator. It does not need to directly modulate the laser at a high-speed frequency.

The subcarrier optical phase-locked loop is shown in Fig. 26.24. The optical phase-locked loop has the advantages of a wide range and fast frequency tuning, and the pseudo-amplitude modulation caused by the frequency change of the tunable laser can also be avoided. The relationship between the maximum linewidth of the laser and the system transmission rate is $\Delta v \leq 5.8 \times 10^{-6} R_b$ [93]. R_b is the bit rate in bit/s, when the bit-error rate is 10^{-10} and the power loss is 1 dB.

Parameters of the Optical Phase-Locked Loop

Following are the parameters of the optical phase-locked loop.

1. Residual phase error: This refers to the phase-error variance between the input and output. Its function is to measure the ability of the optical phase-locked loop system to track the signal light, that is, the phase shift generated when the local oscillator follows the signal light.
2. Tracking performance: In the locked state, the output light phase can stably track the input light phase and the error between the two is small. When the input light phase changes, the output light automatically tracks it within a certain range. The length of the tracking time and size of the steady-state error are both parameters used to evaluate the tracking performance of the system.

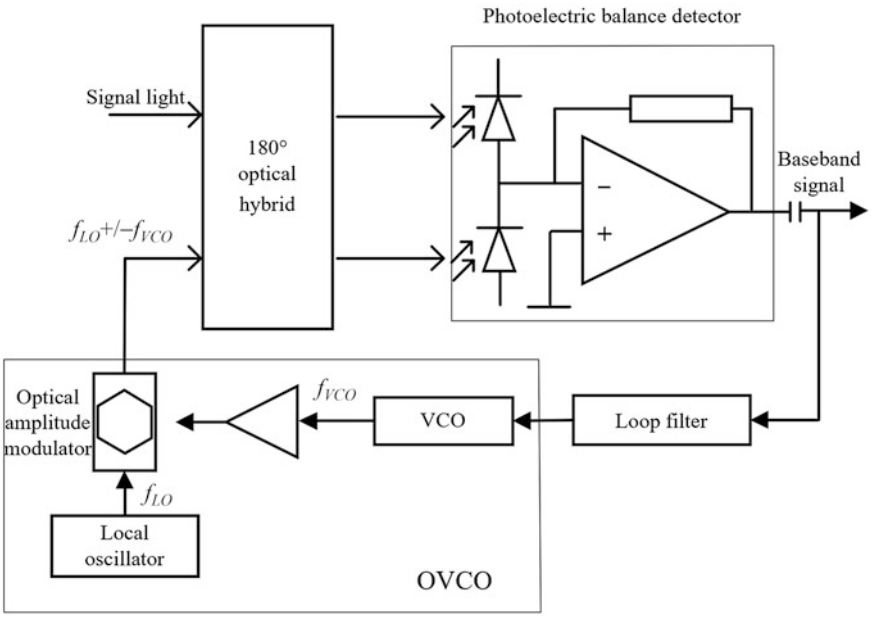


Fig. 26.24 Subcarrier optical phase-locked loop [45]

3. Loop bandwidth: The loop bandwidth refers to the −3-dB bandwidth of the closed-loop transfer function of the optical phase-locked loop, which is the maximum bandwidth at which the local oscillator laser can track the phase of the signal laser. Beyond this range, the local oscillator laser cannot automatically track the signal light. The greater the loop bandwidth, the better the system dynamics.
4. Lock time: This refers to the time required to recover the phase error to 1% when the input light exhibits a step jump. Thus, the locking time should be as short as possible.
5. Laser linewidth: This refers to the full half-peak width of the laser spectrum. This is because the laser is affected by elements of the external environment, such as excited atoms, phase noise, spontaneous ion emission, and temperature jitter. The lower the linewidth, the higher the spectral purity, the better the monochromatic performance of the laser, and the stronger the coherence.

Performance Comparison of Various Optical Phase-Locked Loops

To achieve high-performance coherent-optical communication, it is necessary to comprehensively consider the characteristics of various types of optical phase-locked loops, including the requirements for laser linewidth, types of transmission

Table 26.3 Comparison of performance parameters of various optical phase-locked loops

Type of OPLL	Balanced loop	Costas loop	Decision-driven loop	SyncBit loop	Dither loop	Subcarrier loop
Residual carrier	Yes	No	No	No	No	Yes
90° optical hybrid	No	Yes	Yes	No	No	No
AC coupling link	Not allowed	Allowed	Allowed	Allowed	Allowed	Not allowed
Transparent link	Yes	Yes	No	No	Yes	Yes
Laser linewidth	$\leq 5.8 \times 10^{-6}$	$\leq 5.02 \times 10^{-4}$	$\leq 4.5 \times 10^{-4}$	$\leq 3.1 \times 10^{-4}$	$\leq 2.6 \times 10^{-4}$	$\leq 5.8 \times 10^{-6}$
Data type	Analog, digital	Analog, digital	Digital	Digital	Analog, digital	Analog, digital
Phase-locked power adjustable	No	No	No	No	Yes	No
Characteristic	Laser phase noise tolerance is weak	Laser phase noise tolerance is weak	Only digital signals can be transmitted	Data preprocessing is required	Constant envelope modulation is required	OVCO scheme

signals, and other characteristics. Table 26.3 summarizes the main parameters and compares the performance of various types of optical phase-locked loops.

By comparing several different optical phase-locked structures and analyzing the table and the previous section, it can be seen that the structure of the balanced phase-locked loop is the simplest, and it does not require a 90° optical hybrid. If the energy loss is considered, the power of the 90° optical hybrid is 3 dB lower than that of the 180° hybrid in theory. The SyncBit optical phase-locked loop must preprocess the transmitted data signal, thus introducing additional bits and resulting in reduced spectral efficiency.

The phase-error variance of the decision-driven phase-locked loop is low. A dither phase-locked loop requires constant envelope modulation and a high laser linewidth. The Costas loop has excellent stability and strong capacity to accommodate laser phase noise, with the best comprehensive performance and wide application. The maximum allowable linewidth of the optical phase-locked loop is related to the communication rate of the system, that is, the higher the communication rate, the narrower the laser linewidth.

The Costas loop has the lowest laser-linewidth requirement. The subcarrier optical phase-locked loop can be seen as adding an optical voltage-controlled oscillator to the balanced optical phase-locked loop; therefore, the requirements for the laser linewidth of the subcarrier optical phase-locked loop are also higher.

Optical Phase-Locked Loops in Optical Wireless Coherent Communication

Optical Phase-Locked Loop for Coherent Optical Communication

To meet the requirements of a wider transmission bandwidth, longer transmission distance, and higher sensitivity of the optical communication system, optical coherent modulation technology is introduced into free-space optical communication to form coherent optical communication technology, which has a greater channel capacity and better frequency selectivity.

Figure 26.25 shows a schematic diagram of a optical wireless coherent communication system; its basic principle is the communication modulation part at the transmitter. The frequency and phase of the optical carrier to be transmitted are modulated according to the signal to be transmitted. After reaching the receiving end through an atmospheric transmission, the signal light and the local oscillator light are mixed in the optical hybrid to generate the differential frequency signal carrying the transmission information. The intermediate-frequency signal is obtained through a photoelectric detector, and the baseband signal can be directly obtained by zero-difference coherent detection.

However, owing to the complexity of the space environment, the frequency fluctuation of the signal light and the local oscillator light will negatively affect the performance of the coherent optical communication system, thus, reducing its advantages. Coherent optical communication based on an optical phase-locked loop can effectively capture the frequency of the signal light and maintain stable tracking. Its main application is to meet high requirements for sensitivity and communication speed; however, it cannot use an optical relay. Areas with high requirements for system power consumption and complexity include high-speed optical transmission networks and inter-satellite optical communication systems.

In a coherent optical communication system, an optical phase-locked loop is used to track the optical phase of the receiver, which to a certain extent, ensures a strict match between the signal light and the local oscillator light when mixing, to ensure coherent signal demodulation.

Figure 26.26 shows a schematic diagram of the optical phase-locked loop used in the optical wireless coherent detection process. Homodyne coherent detection has

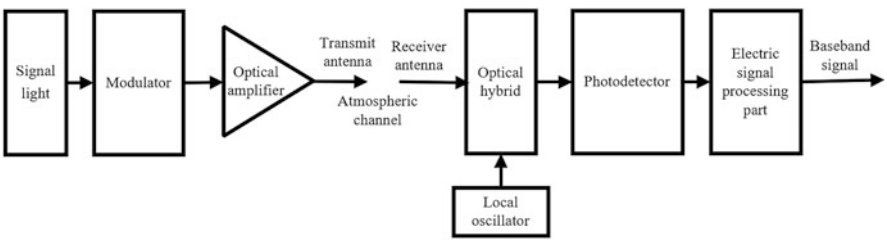


Fig. 26.25 Optical wireless coherent communication system

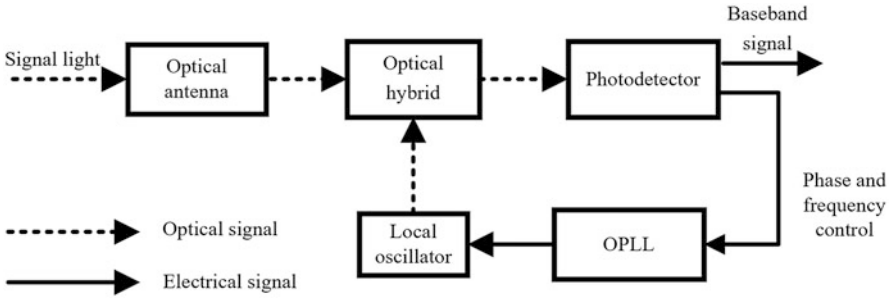


Fig. 26.26 Optical phase-locked loop in a coherent optical communication system

the advantage of directly obtaining the baseband signal [117]. An optical phase-locked loop is crucial for coherent optical communication systems.

Optical phase-locked loop technology receives and demodulates signals coherently in the optical domain and directly recovers the data signals. When the natural frequency difference between the input signal light and the local oscillator light is within a certain range, the optical phase-locked loop can ensure that the difference between the phases of the local oscillator optical signal and the input optical signal is always maintained in a fixed and limited range. This process, controlled by an optical phase-locked loop, is called the capture process.

If the phase difference between the final input optical signal and the local oscillator optical signal is stable, it is called the synchronization state or phase-locked state. In contrast, when the phase difference increases with time, it enters the lock-out state.

The working principle of homodyne coherent detection is shown in Fig. 26.26. The signal light and local oscillator light received by the receiving antenna are first sent into the optical hybrid for coherent mixing. Then, the optical-detection circuit is used to convert the obtained intermediate-frequency signal into the corresponding electrical signal, and then the phase-difference information is fed back to the optical phase-locked loop. This realizes the fast frequency tuning, optical-phase synchronization, and signal demodulation of the local oscillator.

Typical Applications of Optical Phase-Locked Loops

A homodyne coherent detection system based on an optical phase-locked loop has a strong anti-interference ability and a longer transmission distance, and is compatible with DPSK, QPSK, and other modulation formats. With the continuous research on homodyne coherent detection based on optical phase-locked loops, various countries have carried out space optical-communication plans and achieved certain results.

Table 26.4 lists the typical applications of optical phase-locked loops in optical wireless coherent communication. The LEO–LEO inter-satellite communication between the US NFIRE satellite and German TerraSAR-X was the first inter-satellite

Table 26.4 Typical applications of optical phase-locked loops in optical wireless coherent communication

Year	Country	Working wavelength	Communication rate	Modulation format	Receiving method	Experimental condition
1989	Germany SOLACOS [25]	1064 nm	600 Mbit/s	BPSK	Homodyne coherent demodulation of Costas loop	Ground terminal analog LEO–GEO
1996	Europe SROIL [37]	1064 nm	1.5 Gbit/s	BPSK	Homodyne coherent demodulation of SyncBit loop	Ground terminal simulation LEO inter-satellite
2008	US NFIRE satellite and German TerraSAR-X [50]	1064 nm	5.625 Gbit/s	BPSK	Homodyne coherent demodulation of Costas loop	LEO–LEO inter-satellite
2011	Japan JAXA [64]	1064 nm	2.5 Gbit/s	BPSK	Homodyne coherent demodulation of Costas loop	Ground terminal analog GEO–LEO
2013	Japan JAXA [73]	1064 nm	7.2 Gbit/s	BPSK	Homodyne coherent demodulation of Costas loop	Ground-side simulation GEO–LEO
2017	University of Electronic Science and Technology of China [78, 84]	1064 nm	1.25 Gbit/s	BPSK	Homodyne coherent demodulation of balanced loop	Ground-side simulation
2018	Chinese Academy of Sciences [80, 91]	1550 nm	1 Gbit/s	BPSK	Homodyne coherent demodulation of Costas loop	Ground-side simulation

coherent laser communication experiment in the world, which is of great significance for the coherent communication of wireless light.

It is predicted that with the demand for continuous communication development toward high capacity, ultra-long distances, and high speed, zero-difference coherent optical communication technology based on optical phase-locked loops is still the ideal choice for countries to realize space optical transmissions.

In addition, optical phase-locked loops play an important role in optical-frequency modulation, continuous-wave measurement, high-stability time-delay simulation for earth gravity-field measurement, large dynamic optical-fiber transmission links, and precision measurements, such as spectral analysis and other fields.

Summary

Both electric phase-locked loops and optical phase-locked loops have the advantages of a stable output-signal frequency, synchronous tracking, high-speed real-time reception, anti-interference, and noise-suppression performance and have important research value. With the development of optoelectronic technology and the continuous improvement of device technology, electric phase-locked loops and optical phase-locked loops have broad application prospects in many emerging research fields. To improve the practical application of phase-locked loop technology, its performance can be further improved and optimized, as summarized below.

1. Further study the relationship between the mixing efficiency and phase-locked loop error, loop bandwidth, and communication error rate to comprehensively improve the comprehensive performance of the phase-locked loop.
2. Improve the frequency-discrimination acquisition circuit. Optimize the acquisition algorithm. Improve the accuracy of the frequency-difference control in the phase-locked loop. Reduce the acquisition time of the loop frequency to quickly lock the phase-locked loop and increase the acquisition range of the loop frequency [98].
3. Expand the automatic loop-gain adjustment function of the optical phase-locked loop. Further, improve the system robustness, control accuracy, and system-communication performance [95].
4. Optimize the parameters and circuit design of the loop filter to enable the optical phase-locked loop to achieve high-precision performance. Improve the locking ability of weak light and more accurately eliminate the self-excited oscillation signal in the circuit [92].
5. Optimize the loop delay in the system and increase the loop bandwidth to ensure that the phase-locked loop can effectively suppress broadband phase noise while ensuring a certain phase margin.
6. Optimize the temperature-control system of the local oscillator and the signal light. Improve the temperature-control accuracy. Stabilize the temperature of the laser resonator and reduce the frequency drift [98].
7. Develop a fully integrated optical phase-locked loop to reduce loop delay. Further, improve the noise performance and dynamic characteristics of the optical phase-locked loop and reduce the volume and weight of the entire system [77].

References

1. Liu Y. Design and implementation of a high-precision phase-locked loop based on sigma-delta modulation [D]. Chengdu: University of Electronic Science and Technology of China, 2018.
2. Wang Y X, Yan S H, Huo M, et al. Phase locking and homodyne detection of repetitive laser pulses [J]. Optics Express, 2020, 28(24): 35588–35601.
3. Leeb W R, Philipp H K, Scholtz A L, et al. Frequency synchronization and phase locking of CO_2 lasers [J]. Applied Physics Letters, 1982, 41(7):592–594.

4. Yang S J, Ke X Z. Carrier frequency stability control in coherent optical communication [J]. Laser & Optoelectronics Progress, 2018, 55(4):60–66.
5. Ke X Z, Wu J L. Principle and application of optical wireless coherent communication [M]. Beijing: Science Press, 2009.
6. Razavi B. A.2-GHz phase-locked loop [J]. IEEE Journal of Solid-State Circuits, 1997, 32(5): 730–735.
7. Bao H Q. Research and Design of Multi-phase Digital Delay Phase-locked Loop [D]. Xi'an: Xi'an University, 2010.
8. Agiwal M, Roy A, Saxena N. Next Generation 5G Wireless Networks: A Comprehensive Survey [J]. IEEE Communications Surveys & Tutorials, 2016, 18(3): 1617–1655.
9. Peng J. Design of all-digital phase-locked loop in system-on-chip [D]. Nanjing: Southeast University, 2006.
10. Best R E. Phase locked loops: design, simulation, and applications [M]. Beijing: Tsinghua University Press, 2004.
11. Xu Y. Research and Design of Low Noise PLL Frequency Synthesizer [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2018.
12. Zhu L J. Research and Design of Bandwidth Adaptive High-order All-digital Phase-locked Loop [D]. Hengyang: University of South China, 2010.
13. Yang S Z, Li X C, Song J G, et al. Fundamentals of phase-locked technology [M]. Beijing: The People's Posts and Telecommunications Press, 1978.
14. Sheng Z. Research and design of fast adaptive all-digital phase-locked loop [D]. Hengyang: University of South China, 2014.
15. Hou Q, Jie C, Yao Y F, et al. Design of a fast all-digital phase-locked loop combined with high-resolution TDC [J]. Journal of Harbin Institute of Technology, 2018, 50(11):83–88.
16. Long H. Design of PLL frequency synthesizer based on CMOS technology [D]. Changsha: Hunan University, 2007.
17. Gardner F. Charge-pump phase-lock loops [J]. IEEE Transactions on Communications, 1980, 28 (11): 1849–1858.
18. Johnson M G, Hudson E L. A variable delay line PLL for CPU-coprocessor synchronization [J]. IEEE Journal of Solid-State Circuits, 1988, 23(5):1218–1227.
19. Riley T A D, Copeland M A. Delta-sigma modulation in fractional-N frequency synthesis [J]. IEEE Journal of Solid-State Circuits, 1993, 28(5):553–559.
20. Enloe L H, Rodda J L. Laser phase-locked loop [J]. Proceedings of the IEEE, 1965, 53(2): 165–166.
21. Mengali U, Pezzani E. Tracking properties of phase-locked loops in optical communication systems [J]. IEEE Trans. Commun, 1978, 26(12): 1811–1818.
22. Steele R C. Optical phase-locked loop using semiconductor laser diodes [J]. Electron Lett, 1983, 19(2):69–71.
23. Kazovsky L G. Decision-driven phase-locked loop for optical homodyne receivers: performance analysis and laser linewidth requirements [J]. Light-wave Technol, 1985, 3(6): 1238–1247.
24. Kazovsky L G. Performance analysis and laser linewidth requirements for optical PSK heterodyne communications systems [J]. Light-wave Technol, 1986, 4(4):415–429.
25. Pribil K, Krieger J, Meier C. High data rate inter-satellite-communication system SOLACOS [J]. Space Communications, 1998, 15(2):97–104.
26. Zhu Z Z. Research on homodyne coherent optical reception technology [D]. Shanghai: Shanghai Jiao Tong University, 2018.
27. Kazovsky L G, Atlas D A. A 1320-nm experimental optical phase-locked loop: performance investigation and PSK homodyne experiments at 140Mb/s and 2Gb/s [J]. Light-wave Technol, 1990, 8(9):1414–1429.
28. Kahn J M, Gnauck A H, Veselka J J, et al. 4-Gb/s PSK homodyne transmission system using phase-locked semiconductor lasers [J]. IEEE Photonics Technology Letters, 1990, 2(4):285–287.
29. Schopflin A, Kugelmeier S, Gottwald E, et al. PSK optical homodyne system with nonlinear phase-locked loop [J]. Electronics Letters, 1990, 26(6):395–396.

30. Norimatsu S, Iwashita K, Sato K. PSK optical homodyne detection using external cavity laser diodes in Costas loop [J]. IEEE Photonics Technology Letters, 1990, 2(5):374–376.
31. Norimatsu S, Iwashita K, Noguchi K.10Gbit/s optical PSK homodyne transmission experiments using external cavity DFB LDs [J]. Electron Lett, 1990, 26(10):648–649.
32. Wandernoth B. 1064 nm, 565 Mbit/s PSK transmission experiment with homodyne receiver using synchronisation bits [J]. Electronics Letters, 1991, 27(19): 1692–1693.
33. Bopp M, Huether G, Spatscheck T, et al. BPSK Homodyne and DPSK Heterodyne Receivers for Free-Space Communication with Nd: Host Lasers [C]//Optical Space Communication. SPIE, Munich, Germany, 1991:199–209.
34. Hornbachner D F, Schreiblehner M A, Leeb W R, et al. Experimental determination of power penalty contributions in an optical Costas-type phase-locked loop receiver [C]//Free-Space Laser Communication Technologies. SPIE, Los Angeles, CA, 1992, 1635:10–18.
35. Huang C, Wang L. Optimum phase deviations between mark and space-state bits in long-haul communication systems with balanced PLL receivers [J]. Light-wave Technol, 1995, 13(10): 1963–1970.
36. Chiou Y, Wang L. Effect of optical amplifier noise on laser linewidth requirements in long haul optical fiber communication systems with Costas PLL receivers [J]. Light-wave Technol, 1996, 14(10):2126–2134.
37. Lutz H. Esas activities towards a short-range laser communications terminal [J]. Space Communications, 1998, 15(2): 71–78.
38. Djordjevic I B, Stefanovic M C. Performance of optical heterodyne PSK systems with Costas loop in multichannel environment for nonlinear second-order PLL model [J]. Journal of Lightwave Technology, 1999, 17(12): 2470–2479.
39. Kudielka K H, Klaus W. Optical homodyne PSK receiver: Phase synchronization by maximizing baseband signal power [C]//IEEE Lasers and Electro-Optics Society. IEEE, 1999, 1: 295–296.
40. Liu L, Xiao J N, Tang J, et al. Experimental Research on Timing Recovery of a Differential Quadrature Phase Shift Keying Coherent Optical Transmission System Based on an Improved Gardner Algorithm [J]. Acta Optica Sinica, 2013, 33(4):23–29.
41. Camatel S, Ferrero V, Gaudino R, et al. Optical phase-locked loop for coherent detection optical receiver [J]. Electronics Letters, 2004, 40(6):384–385.
42. Lange R, Smutny B. Optical inter-satellite links based on homodyne BPSK modulation: Heritage, status and outlook [C]//Free-Space Laser Communication Technologies. SPIE, San Jose, CA, 2005, 5127:1–12.
43. Herzog F, Kudielka K, Erni D, et al. Optical phase locked loop for transparent inter-satellite communications [J]. Optics express, 2005, 13(10): 3816–3824.
44. Zhang R Z, Yang C L. Active phase-locking technology of laser arrays [J]. Semiconductor Optoelectronics, 2005, 26(1):54–57.
45. Camatel S, Ferrero V. Homodyne coherent detection of ASK and PSK signals performed by a subcarrier optical phase-locked loop [J]. Photonics Technology Letters, IEEE, 2006, 18(1): 142–144.
46. Lange R, Smutny B. BPSK laser communication terminals to be verified in space [C]//IEEE Military Communications Conference. IEEE, Monterey, CA, USA, 2004:441–444.
47. Friederichs L, Sterr U, Dallmann D. Vibration Influence on Hit Probability During Beaconless Spatial Acquisition [J]. Journal of Lightwave Technology, 2016, 24(10):2500–2509.
48. Toyoshima M, Takayama Y, Takahashi T, et al. Laser Beam Propagation In Ground-to-oicets Laser Communication Links [J]. The Journal of Space Technology and Science, 2007, 23(2): 30–45.
49. Liang W, Satyan N, Aflatouni F, et al. Coherent beam combining with multilevel optical phase-locked loops [J]. Journal of the Optical Society of America B, 2007, 24(12):2930–2939.
50. Romeiser R, Runge H. Theoretical Evaluation of Several Possible Along-Track InSAR Modes of TerraSAR-X for Ocean Current Measurements [J]. IEEE Transactions on Geoscience and Remote Sensing, 2007, 45(1):21–35.

51. Zhang Y L, Wang Z Y. An optical phase-locked loop clock extraction system based on EAM four-wave mixing effect [J]. Optical communication technology, 2007, (4):43–44.
52. Quan S, Yao M Y, Zhang H M, et al. 40 Gb/s Clock Recovery Based on Optical Phase-Locked Loop [J]. Acta Optica Sinica, 2007, 27(8):1382–1386.
53. Camatel S, Ferrero V. Design, analysis and experimental testing of BPSK homodyne receivers based on subcarrier optical phase-locked loop [J]. Journal of Lightwave Technology, 2008, 26(5): 552–559.
54. Sodnik Z, Furch B, Lutz H. Optical Intersatellite Communication [J]. IEEE journal of selected topics in quantum electronics, 2010, 16 (5): 1051–1057.
55. Ristic S, Bhardwaj A, Rodwell M J, et al. An optical phase-locked loop photonic integrated circuit [J]. Journal of Lightwave Technology, 2010, 28(4): 526–538.
56. Heine F, Kampfner H, Lange R, et al. Laser communication applied for EDRS, the European data relay system [J]. CEAS Space Journal, 2011, 2(4): 1–6.
57. Xu Q, Mondragón A A, Gallion P, et al. Homodyne in-phase and quadrature detection of weak coherent states with carrier phase tracking [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2009, 15(6): 1581–1590.
58. Wang H. Design of homodyne BPSK system for coherent optical communication [D]. Chengdu: University of Electronic Science and Technology of China, 2009.
59. Ponnampalam L, Steed R J, Fice M J, et al. A compact tunable coherent terahertz source based on an hybrid integrated optical phase-lock loop [C]//Microwave Photonics (MWP), 2010 IEEE Topical Meeting on. IEEE, Montreal, CA, 2010: 151–154.
60. Satyan N, Sendowski J, Vasilyev A, et al. Phase noise reduction of a semiconductor laser in a composite optical phase-locked loop [J]. Optical Engineering, 2010, 49(12):1127–1134.
61. Sakamoto T, Chiba A, Kanno A, et al. Real-time homodyne reception of 40-Gb/s BPSK signal by digital optical phase-locked loop [C]//European Conference on Optical Communication. IEEE, Torino, Italy, 2010:1046–1048.
62. Du X. Research on phase synchronization and correction technology in coherent optical communication system [D]. Beijing: Beijing University of Posts and Telecommunications, 2010.
63. Shoji Y, Takayama Y, Toyoshima M, et al. Feasibility study of coherent LEO-ground link system using an optical injection phase lock loop technique [C]// International Conference on Space Optical Systems and Applications. IEEE, Santa Monica, CA, 2011:318–324.
64. Ando T, Haraguchi E, Tajima K, et al. Homodyne BPSK receiver with Doppler shift compensation for inter satellite optical communication [C]//Space Optical Systems and Applications. IEEE, Santa Monica, CA, 2011:279–285.
65. Steed R J, Pozzi F, Fice M J, et al. Monolithically integrated heterodyne optical phase-lock loop with RF XOR phase detector [J]. Optics Express, 2011, 19 (21): 20048–20053.
66. Qiu Z H. Study of Costas optical phase-locked loop [D]. Chengdu: University of Electronic Science and Technology of China, 2011.
67. Ding Z. Design and simulation of optical phase-locked loop [D]. Chengdu: University of Electronic Science and Technology of China, 2011.
68. Park H C, Lu M, Bloch E, et al. 40Gbit/s coherent optical receiver using a Costas loop [J]. Optics Express, 2012, 20(26):197–203.
69. Lu M, Park H, Bloch E, et al. Highly integrated optical heterodyne phase-locked loop with phase/frequency detection [J]. Optics express, 2012, 20(9): 9736–9741.
70. Mizutori A, Sugamoto M, Koga M.12.5-Gbit/s BPSK Stable Optical Homodyne Detection Using 3-kHz Spectral Linewidth External-Cavity Laser Diode [C]// European Conference and Exhibition on Optical Communications. IEEE, Amsterdam, 2012:1046–1048.
71. Xu Z X. Automatic control of laser frequency and process in cold atom interference experiment [D]. Hangzhou: Zhejiang University, 2012.
72. Peng J B, Bian Z L, Hao Y Q, et al. Optical frequency locking system for Brillouin distributed optical fiber sensor [J]. Chinese Journal of Lasers, 2012, 39(7):126–129.
73. Ando T, Haraguchi E, Tajima K, et al. Optical homodyne BPSK receiver with Doppler shift compensation for LEO-GEO optical communication [C]//Lasers and Electro-Optics Pacific Rim. IEEE, Kyoto, Japan, 2013:279–285.

74. Ding X, Sang L, Chen Z, et al. High precision 6.8GHz phase locking of coherent laser beams for optical lattice experiment [C]//Optical Systems and Modern Optoelectronic Instruments. International Conference on Optical Instruments and Technology, Beijing, 2013:1–7.
75. Meng Z M, Zhang J. Phase-Locked Laser System Used for Electromagnetically Induced Transparency Spectral Experiments [J]. Acta Optica Sinica, 2013, 33(7):219–228.
76. Balakier K, Fice M J, Ponnampalam L, et al. Monolithically Integrated Optical Phase Lock Loop for Microwave Photonics [J]. Lightwave Technology Journal of, 2014, 32(20): 3893–3900.
77. Shi H X. Research on Optical Phase-Locked Loop [D]. Shanghai: Shanghai Jiao Tong University, 2015.
78. Yue H. Research on homodyne coherent optical communication technology based on 90° optical mixer [D]. Chengdu: University of Electronic Science and Technology of China, 2014.
79. Koga M, Shigeta Y, Shirazawa F, et al. Costas Loop Homodyne Detection for 20-Gb/s QPSK Signal on the Optical Frequency Synchronous Network [J]. Journal of Lightwave Technology, 2015, 33(23):4752–4760.
80. Zhang Z, Sunj F, Lu B, et al. Costas Optical Phase Lock Loop System Design in Inter-Orbit Coherent Laser Communication [J]. Chinese Journal of Lasers, 2015, 42(08):177–182.
81. Jiang X Z. Design and preliminary experimental test of heterodyne weak light phase-locking scheme [D]. Wuhan: Huazhong University of Science and Technology, 2015.
82. Wang Y, Li B, Guo Y, et al. Tracking performance of optical phase locking loop with frequency discrimination and control subloop [J]. Optical Engineering, 2016, 55(5): 1–7.
83. Ma Z J, Wang D, Wu J Z, et al. Design of Optical Phase Locked Loop for Atoms-Lights Coupling Systems [J]. Journal of Quantum Optics, 2016, 22(2):193–198.
84. Wang Y, Wang C, Tao Y, et al. Fast Frequency Acquisition and Phase Locking of Nonplanar Ring Oscillators [J]. Applied Sciences, 2017, 7(10): 1032.
85. Chang S, Tong S F, Jiang H L, et al. Optical Phase-Locked Loop Technology in Inter-Satellite High-Speed Coherent Laser Communication Systems [J]. Acta Optica Sinica, 2017, 37(2): 76–84.
86. Zhu Z, Zhou H, Xie W, et al. 10-Gb/s homodyne receiver based on Costas loop with enhanced dynamic performance [C]//International Conference on Optical Communications and Networks. IEEE, Wuzhen, 2017:1184–1186.
87. Zhou H, Zhu Z, Xie W, et al. Investigation of signal power splitting ratio for BPSK homodyne receiver with an optical Costas loop [J]. Optical Engineering, 2018.57(8):1–7.
88. Ashtiani F, Sanjari P, Idjadi M H, et al. High-resolution optical frequency synthesis using an integrated electro-optical phase-locked loop [J]. IEEE Transactions on Microwave Theory and Techniques, 2018, 66(12): 5922–5932.
89. Wang J, Chen D J, Cai H W, et al. Optical Phase-Locked Loop of Single Section Distributed Feedback Semiconductor Laser [J]. Chinese Journal of Lasers, 2018, 45(4):29–34.
90. Liu X, Li J, Guo S, et al. Study on optical phase-locked loop technology in space homodyne coherent optical communication [J]. Radio Engineering, 2018, 48(8):718–726.
91. Yue C, Li J, Sun J, et al. Homodyne coherent optical receiver for intersatellite communication [J]. Applied optics, 2018, 57(27):7915–7927.
92. Tao Y P. Research on broadband low-delay optical phase-locked loop [D]. Chengdu: University of Electronic Science and Technology of China, 2021.
93. Liu Y. Research on detection technology of inter-satellite homodyne coherent laser communication based on AOFS tuning [D]. Changchun: Changchun University of Science and Technology, 2021.
94. Pan B, Chen H. Design of carrier acquisition loop based on optical phase-locked loop [J]. Information & Communications, 2019, (10):44–45.
95. Wang J Y. Research on Fast Acquisition Technology of Optical Phase Locked Loop in Space Coherent Laser Communication [D]. Changchun: Changchun University Of Science And Technology, 2023.

96. He B W. Research on Homodyne Frequency acquisition Technology Based on Orbit Prediction in Satellite Optical Communication [D]. Harbin: Harbin Institute of Technology, 2023.
97. Wang H. Research on Self-Injection and Optical Phase-Locked Loop Locking of Distributed Feedback Semiconductor Laser [D]. Changchun: Changchun University Of Science And Technology, 2024.
98. Huo M. High-sensitivity homodyne coherent detection of laser pulse echo [D]. Chengdu: University of Electronic Science and Technology of China, 2024.
99. Zhang C C, Wang J B, Yin C, et al. Research on High Stability Laser Frequency Stabilization Method Based on Optical Phase-Locked Loop [J]. Acta Metrologica Sinica, 2022, 43(9): 1154–1160.
100. Fang T P, Hu H, Yan P P, et al. Design and Experimental Verification of Coded Synthetic Aperture Lidar System Based on Optical Phase-Locked Loop [J]. Acta Optica Sinica, 2022, 42(6):136–143.
101. Dou X J. Design and Research of CMOS Phase-locked Loop and Delay Phase-locked Loop [D]. Beijing: Peking University. 2002.
102. Guo F Q. Design of fractional frequency division phase-locked loop based on phase interpolation technology [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2021.
103. Liu J J. Design and Implementation of All-digital Phase-locked Loop Circuit [D]. Chengdu: University of Electronic Science and Technology of China, 2011.
104. Zhang W C. All-digital phase-locked loop technology for wireless communication [D]. Shanghai: Shanghai Jiao Tong University, 2010.
105. Dong Y. Phase-Locked Loop Circuit Design [J]. Computer Knowledge and Technology, 2011, 7(18):4459–4461.
106. Young H S, Seon K L, Jae Y S. A 1-GHz Digital PLL With a 3-ps Resolution Floating-Point-Number TDC in a 0.18-μm CMOS [J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2011, 58(2): 70–73.
107. Yang C N. Research and design of 2.4GHz CMOS all-digital phase-locked loop [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2015.
108. Takashi T, Mitsuji O, Masaki K, et al. A 2.1-to-2.8-GHz Low-Phase-Noise All-Digital Frequency Synthesizer With a Time-Windowed Time-to-Digital Converter [J]. IEEE Journal of Solid-State Circuits, 2010, 45(12): 2582–2590.
109. Shan C H, Meng X Y. The Design of All-Digital Phase-Locked Loop Based on FPGA [J]. Application of Electronic Technique, 2001, 27(9):58–60.
110. Lee S K, Seo Y H, Park H J, et al. A 1 GHz ADPLL With a 1.25ps Minimum-Resolution Sub-Exponent TDC in 0.18 um CMOS [J]. IEEE Journal of Solid-State Circuits, 2010, 45(12): 2874–2881.
111. Murmann B. Digitally Assisted Analog Circuits [J]. IEEE Micro, 2006, 26(2):38–47.
112. Riley T A D, Copeland M A, Kwasniewski T A. Delta-sigma modulation in fractional-N frequency synthesis [J]. IEEE Journal of Solid-State Circuits, 1993, 28(5):553–559.
113. Yu J, Maloberti F. A low-power multi-bit ΔΣ modulator in 90nm digital CMOS without DEM [J]. IEEE Journal of Solid-State Circuits, 2005, 40(12):2428–2436.
114. Gardner F M. Phaselock techniques [J]. Systems, Man and Cybernetics, IEEE Transactions on, 1984, 14(1):170–171.
115. Yu J. Experimental Study on Frequency Modulator 4046 Based on PLL [J]. Experiment Science and Technology, 2013, 11(5):4–6+29.
116. Ke X Z, Yin Z Y. Coding theory in wireless laser communication system [M]. Beijing: Science Press, 2008.
117. Ke X Z, Dong K. Optical Wireless Communication [M]. Beijing: Beijing Springer Press, 2022
118. Kazovsky L G. Balanced phase-locked loops for optical homodyne receivers: performance analysis, design considerations and laser linewidth requirements [J]. Light-waveTechnol, 1986, 4(2):182–195.
119. Glatt W, Schreiblehner M A. Optical PSK homodyne system using a switched residual carrier for phase synchronization [J]. Electronics Letters, 1996, 32(15): 1386–1387.

Deformable Mirrors and Their Control Algorithms

27

Contents

Introduction 988
Foreign Research Progress 988
Domestic Research Progress 991
Deformable Mirrors 993
Classification of Deformable Mirrors 993
Typical Deformable Mirrors 995
Performance Index of Deformable Mirrors 999
Model of a Deformable Mirror 999
Deformable-Mirror Control Algorithms 1000
Application of Deformable-Mirror Control Algorithms in Adaptive-Optics Systems 1010
Summary 1012
References 1012

Abstract

Deformable mirror is the key device to realize wavefront correction in adaptive-optics system, and its performance directly determines the wavefront-distortion correction ability of the system. This chapter summarizes the research progress of deformable mirrors at home and abroad. Taking several typical deformable mirrors as examples, the classification, modeling, and main performance indicators of deformable mirrors are introduced. Mirrors, thin film deformable mirrors, dual-piezoelectric deformable mirrors, micro-electromechanical systems (MEMS), deformable mirrors, and voice-coil deformable mirrors are introduced in detail. The correction accuracy of different control algorithms for wavefront distortion is analyzed, and the work done by Xi'an University of Technology in this field is also introduced. Finally, the future technological breakthroughs and improvement directions in this field are pointed out.

Keywords

Deformable mirror · Adaptive optics · Control algorithm · Research progress

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_27

Introduction

A deformable mirror is a key device for correcting wavefronts in adaptive-optics systems, and its performance directly determines the wavefront-distortion correction ability of the system. This chapter reviews the research progress on deformable mirrors at home and abroad.

Taking several typical deformable mirrors as examples, their classification, modeling, and main performance indicators are introduced. Moreover, different types of deformable mirrors are discussed, such as separation actuator, spliced, thin film, dual piezoelectric, micro-electromechanical system (MEMS), and voice coil. Finally, the technical breakthroughs and directions for future improvement in this field are presented.

Deformable mirrors (DMs) are widely used in various optical systems to correct wavefront distortions. They correct the distorted wavefront by changing the optical path of the light-wave transmission. They offer the advantages of large travel and good linearity. Many types of deformable mirrors exist, and they each have different characteristics, such as different hysteresis, response times, and deformations. To improve the correction accuracy, it is necessary to control the deformable mirror using a control algorithm. Therefore, research on deformable mirror control algorithms has become a key task in adaptive optics.

Foreign Research Progress

In 1953, Babcock proposed the concept of adaptive optics. Its main idea was to place a "corrector" on the pupil surface, and change the surface shape of the corrector through real-time control of the optical wavefront to compensate for distortion caused by atmospheric disturbances. The corrector mentioned by Babcock in his groundbreaking discussion was called an "eidophor." Its working principle involved applying a thin oil film to the surface of a mirror, and applying a control voltage to the oil film. Under the effect of electrostatic forces, the oil film thickness changed with the spatial distribution of the charge, thus affecting the optical path of the incident light [1].

In 1976, Yellin et al. designed and fabricated a thin-film deformable mirror by comprehensively considering the frequency response and driving signal. A thin-film deformable mirror mainly adopts an ultrathin electrostatic film with a thickness of 0.5–1 μm. Because of the advantages of thin films, such as low voltage and zero hysteresis, thin-film deformable mirrors are the best choice for image compensation systems, both physically and in engineering, and have been widely used in the space field [2].

In 1979, Kokorowsky and Schwartz et al. proposed a dual-piezoelectric deformable mirror. It was composed of two piezoelectric-ceramic sheets bonded together with opposite polarization directions. Using the transverse–converse piezoelectric effect of piezoelectric materials, a voltage was applied to the sheet. One sheet

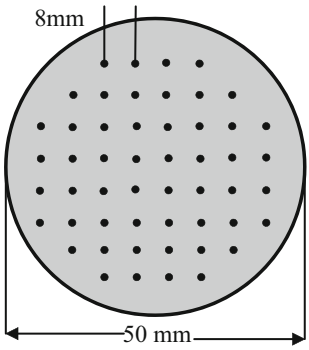


Fig. 27.1 Arrangement of deformation mirror actuators [6]

expanded and the other sheet shrunk in the direction perpendicular to the polarization, which deformed the surface shape of the deformable mirror [3, 4].

In 1982, Freeman et al. developed a phase-correction device that used a deformable mirror with a refrigeration system. The actuator arrangement is shown in Fig. 27.1. The deformable mirror had 52 control units. They used an 8 × 8 square array layout. The distance between adjacent actuators was 8 mm and the effective luminous aperture was 50 mm. A single actuator consisted of 99 piezoelectric chip stacks with a 15% coupling coefficient [5].

In the early 1990s, Hulburd et al. developed a spliced deformable mirror and real-time atmospheric compensation system. Each submirror of the spliced deformable mirror had three actuators made of piezoelectric-ceramic materials [7].

In 1994, Takami et al. designed a thin-film deformable mirror for the adaptive-optical system of the Subaru Telescope in Japan. The thin film of the deformable mirror was a 2-μm-thick nitrocellulose material. The diameter of the deformable mirror was 50.8 mm, the effective luminous aperture was 25 mm, and the surface shape had a 0.03λ root-mean-square (RMS) error. The electrode actuator was a 5 × 5 square array and the distance between adjacent actuators was 4 mm [8].

In 1997, Oppenheimer et al. developed a deformable mirror. The number of actuator control units was 349, and the mirror surface was 2-mm-thick silver plating. After flattening, the surface shape was better than 19-nm RMS [9].

In 1997, Bifano et al. developed a continuous-surface deformable mirror using micromachining technology. The deformable mirror had nine actuators with a 2-μm spacing. A performance test showed that the influence range of the adjacent actuators was 5–10% [10].

In 1998, Lowrey et al. tested the performance of three deformable mirrors with 349, 577, and 941 elements. The actuator spacing of the three deformable mirrors was 7 mm. The effective luminous aperture of the 349-element deformable mirror was 148 mm. After flattening, the surface error was 20-nm RMS. The diameter of the 577-unit deformable mirror was 182 mm, and the effective luminous aperture was 190 mm. After flattening, the surface error was 19-nm RMS.

The 941-unit deformable mirror had 941 actuators in a 35 × 35 square-array arrangement. The working voltage was ±30 V, the bias voltage was 70 V, the

maximum travel was ±2 μm, and the effective luminous aperture of the deformable mirror was 288 mm [11]. The 941-element deformable mirror was applied to the adaptive-optical system of a SOR3.5 telescope to compensate for the phase interference in the atmosphere [12].

In 2000, Paterson et al. developed a low-cost adaptive-optical system using a 37-element thin-film deformable mirror. The film was made of silicon nitride with a 500- to 700-nm thickness. The effective aperture of the film-deformable mirror was 15 mm, distance between adjacent actuators was 1.75 mm, and maximum deformation was 1000 nm [13]. Thin-film deformable mirrors can measure and correct human-eye aberrations in real time and are widely used in human-eye adaptive-optics technology [14].

To satisfy the requirements of space-coronal observation telescopes and planetary high-resolution imaging telescopes [15], Ealey et al. developed a deformable mirror with a high-density actuator array in 2004. This deformable-mirror technology was based on a modular actuator array and could be expanded to thousands of actuator elements. The number of actuator control units was 1946. They were arranged in a 42 × 42 square array, and the spacing between adjacent actuators was 1 mm [16].

In 2005, Sinquin et al. developed a 188-electrode deformable mirror for the Subaru Telescope and successfully applied it to the adaptive-optics system, enabling it to reach the diffraction limit [18]. The performance of the deformable mirror was tested and is described in detail in [17].

In 2008, Sinquin et al. developed an actuator module for a high-density deformable mirror with the support of the European Southern Observatory. It had a 50 × 50 square array with a 1-mm distance between adjacent actuators. They predicted that with the development of technology, high-density deformable mirrors could greatly increase, with 30,000 actuator control units, a 200-mm maximum effective luminous aperture, and 3-μm travel [18].

In 2008, Comelissen et al. developed a 4096-unit MEMS deformable mirror to correct higher-order wavefront aberrations. The effective aperture size of the deformable mirror was 27.2 mm, and the actuators were in a 64 × 64 square arrangement with a 400-μm actuator spacing. The working bandwidth was 2.5 kHz, and the travel could reach 4 μm [18].

In 2010, Dekany et al. developed a 4356-element deformable mirror for a Palomar 5.1-m telescope, enabling the ultraprecision near-infrared observation of exoplanets [19].

In 2012, Verpoort et al. developed a new single-pressure deformable mirror for real-time intracavity phase control of high-power continuous lasers and introduced its dynamic characteristics, influence function, and other characteristics. The effective aperture of the deformable mirror was 9 mm. In the low-order aberration mode, it produced several micrometers of travel [20].

In 2016, Kuiper et al. designed a 61-element electromagnetic deformable mirror. Its diameter was 200 mm, the actuator coupling efficiency was 24%, and the distance between adjacent actuators was 22 mm. Deformable mirrors are used in various applications in terrestrial astronomy and space fields [21]. Two years later, Kuiper et al. studied an actuator based on the variable reluctance principle, developed a

deformable mirror using this technology, and tested its performance. The surface diameter of the deformable mirror was 160 mm, with 57 actuator control units [22].

Domestic Research Progress

In 1995, Jiang et al. developed a system to correct wavefront distortion under different turbulence intensities. A separation-actuator continuous-surface deformable mirror was used to correct higher-order aberrations in the system. It had 55 deformable-mirror actuator control units. The effective luminous aperture was 100 mm, the deformation was ±1.5 μm, nonlinear hysteresis was less than ±4%, and resonant frequency was greater than 2 kHz [23].

In 1998, Jiang et al. also developed a 61-unit adaptive-optical system and docked it with a tracking and pointing system, realizing the first atmospheric-transmission compensation experiment in China. The effective aperture of the deformable mirror was 120 mm, the maximum deformation was ±3 μm [a lot], the nonlinear hysteresis was less than 5%, the resonant frequency was greater than 2 kHz, and the interconnection value was 8–12% [24].

In 1998, Ling et al. developed a discrete piezoelectric deformable mirror with 61 actuation units and applied it to an adaptive-optical system. They successfully completed the docking between acquisition, pointing, and tracking (APT) and AO systems, and improved the APT tracking accuracy [25].

In 2006, Xiang et al. developed a new deformable mirror with a transparent electrode using MEMS technology and tested its performance. A schematic diagram is shown in Fig. 27.2. The deformable mirror could perform bidirectional deformations. When a certain driving voltage was applied, the maximum deformation range of the deformable mirror reached 19 μm [26].

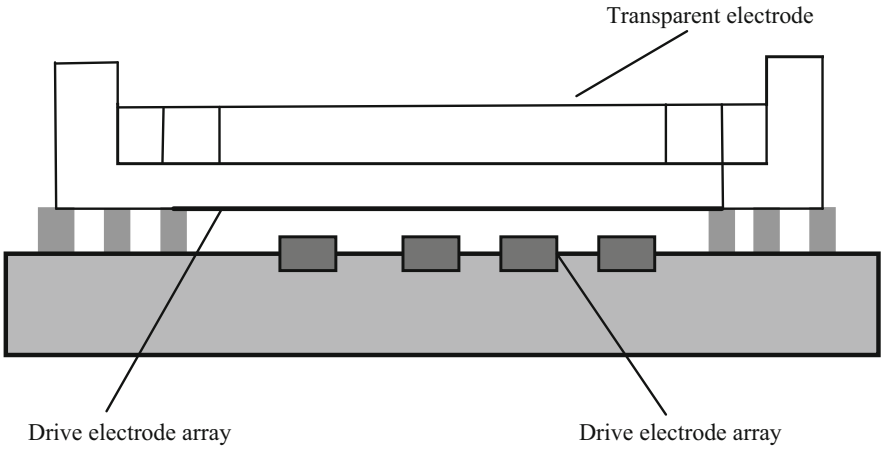


Fig. 27.2 Schematic diagram of a deformable mirror with transparent electrodes [26]

In 2008, Zhou et al. developed a 20-element dual-piezoelectric deformable mirror, and tested and analyzed the performance. The results showed that the dual-piezoelectric deformable mirror was suitable for low-order aberration corrections [27].

In 2010, Wei et al. developed a 127-element deformable mirror and applied it to a 1.8-m astronomical telescope to achieve high-resolution star imaging [28].

In 2012, Lin et al. initially developed the DM21-1 deformable mirror. The effective luminous aperture of the deformable mirror was 50 mm. The actuator was composed of a piezoelectric-ceramic material. It had 21 actuator control units, arranged in a 5 × 5 square array. The distance between adjacent actuators was 9 mm. Figure 27.3 shows the actuator arrangement of the deformable mirror [29].

In 2013, Lin et al. developed a 97-element piezoelectric deformable mirror. The effective luminous aperture of the deformable mirror was 70 mm, the actuators were arranged in an 11 × 11 square array, the distance between adjacent actuators was 6.5 mm, and the intersection value was 4.8% [30].

In the same year, Lin et al. developed a deformable mirror that separated the continuous surface of the actuator. The effective luminous aperture was 90 mm, and the number of actuator control units was 137, arranged in a 13 × 13 square array. The distance between adjacent actuators was 7 mm, the interconnection value was 22–26%, and the maximum deformation was ±2.5 μm. The nonlinear hysteresis was less than 5%, the total resonant frequency was greater than 12 kHz, and the surface shape was better than $\lambda/50(\lambda = 632.8$ nm) RMS after active flattening [31].

To ensure the normal use of adaptive-optics systems at low temperatures and correct higher-order aberrations, You et al. developed a 913-element continuous-surface deformable mirror in 2019 and tested and analyzed its main performance indicators. The results showed that the maximum deformation of a single actuator of the deformable mirror was ±3.5 μm. The interaction value of adjacent actuators was 9.3% [32].

In 2022, Du et al. studied the submirror assembly of a spliced deformable mirror and developed an integrated design to ensure the surface accuracy and stability of the

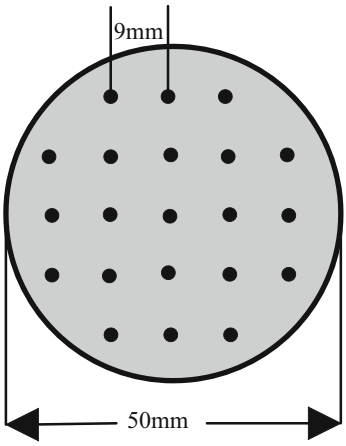


Fig. 27.3 Actuator arrangement of a 21-unit deformable mirror [29]

primary mirror. The spliced deformable mirror was composed of three hexagonal Zerodur submirrors. After integrating the design, the total mass of the submirror was 1.74 kg, and the surface shape of the submirror was better than 0.019λ RMS. It met the design requirements and provided a new idea for the integrated design of space mirrors [33].

Deformable Mirrors

A deformable mirror compensates for wavefront aberrations in a light wave by changing the light wave's optical-path difference. It contains many units. Each independent unit has its own actuator. When a driving voltage is applied to the actuator, the surface shape of the deformable mirror changes. Deformable mirrors play an important role in adaptive-optics systems. Therefore, the study of deformable mirrors affects the development of the entire system.

Classification of Deformable Mirrors

Deformable mirrors are widely used in the adaptive-optics field. One method classifies deformable mirrors according to their surface shape, and the other classifies them according to their driving mode. In this chapter, two classification bases are introduced in detail. The deformable mirrors used in adaptive-optics applications by Ke Xizheng et al. of Xi'an University of Technology are introduced.

Divided by Surface Shape

Depending on their surface shape, deformable mirrors can be divided into two types: discrete surface (whose surface structure is shown in Fig. 27.4) and continuous surface (whose surface structure is shown in Fig. 27.5) [34].

Most early deformable mirrors were discrete-surface deformable mirrors composed of several independent submirrors. Each independent mirror modified a wavefront by controlling the translational or tilting motion of the actuator of each independent submirror. This type of deformable mirror is easy to control and assemble; however, its surface is discrete and large. Gaps between adjacent mirror

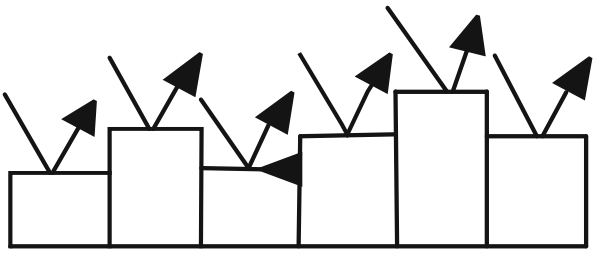


Fig. 27.4 Surface structure of a discrete-surface deformable mirror [35]

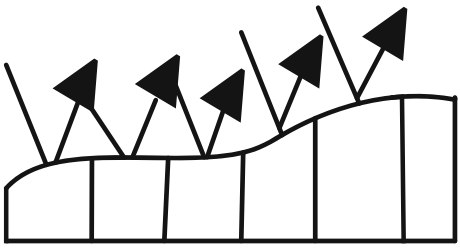


Fig. 27.5 Surface structure of a continuous-surface deformable mirror [35]

surfaces lead to a loss of light energy and affect the correction accuracy of the deformable mirror.

A continuous-surface deformable mirror (CSDM) moves the surface of the deformable mirror through different driving modes to change its shape. This type of deformable mirror has a continuous surface without gaps; therefore, no light is lost and the phase is continuous. Consequently, continuous-surface deformable mirrors have a higher correction accuracy than discrete-surface deformable mirrors [35].

Divided by Driving Mode

According to their driving mode, deformable mirrors can be divided into piezoelectric, electromagnetic, and electrostrictive [36]. The piezoelectric deformable mirror has a hysteresis; therefore, a specific control algorithm must be used to eliminate the hysteresis to ensure the correction accuracy; however, this increases the complexity of the correction algorithm and reduces the real-time performance of the system [34].

The electromagnetic deformable mirror is mainly composed of a coil and a magnet. When the coil is energized, the interaction force between the coil and magnet causes the mirror surface to move up and down, changing its shape. Electromagnetic deformable mirrors have large travel and wide correction ranges, and are widely used in optical communication systems [35].

An electrostrictive deformable mirror is also called a lead–magnesium–niobate (PMN)-deformable mirror. The actuator is made of electrostrictive materials (Pb (Mg, Nb,) O_3 (PMN)). The main difference between piezoelectric deformable mirrors and electromagnetic deformable mirrors lies in the different materials used to make actuators.

Ke Xizheng et al. at Xi'an University of Technology have been successful in adaptive-optics applications. The deformable mirror used in this application was the DM69, which is a continuous deformable mirror driven by electromagnetism. The DM69 is controlled by applying a control voltage to the 69 actuators in the deformable mirror. The effective luminous aperture is 10.5 mm, the weight is 0.45 kg, the distance between adjacent actuators is 1.5 mm, and the maximum correction amount of the tilt wavefront peak–valley (PV) is 60 μm. The working temperature is −10–35 °C, and the nonlinear error is less than 3%. A structural diagram of the DM69 deformable mirror is shown in Fig. 27.6. Figure 27.7 shows the layout and sequencing diagram of the DM69 deformable-mirror actuator.

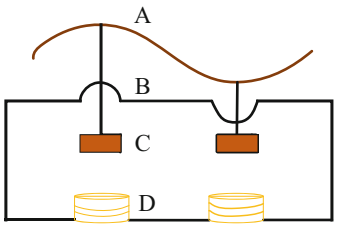


Fig. 27.6 Structural diagram of a deformable mirror

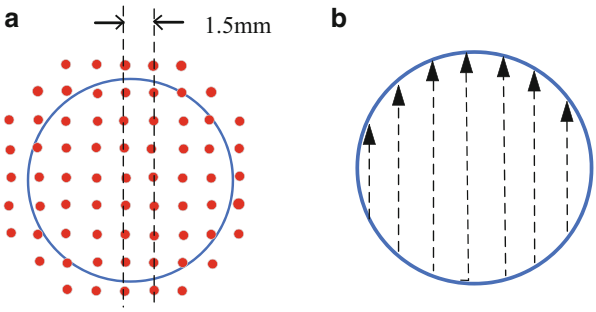


Fig. 27.7 (**a**) Arrangement and (**b**) sequence diagram of a deformable-mirror actuator

Typical Deformable Mirrors

As key wavefront-correction components, deformable mirrors have been studied for many years, and researchers have summarized their rich experience. The following are typical deformable mirrors, according to their structure and working principle.

Separation-Actuator Continuous-Surface Deformable Mirror

The structure of a continuous-surface deformable mirror with separate actuators is shown in Fig. 27.8. From bottom to top, the deformable mirror is composed of three parts: the base, actuator, and specular surface (mirror). The reflector adopts a continuous thin lens, and the surface shape of the thin lens changes with the actuator displacement. When the actuator moves, the base does not deform, but the thin mirror connected to the actuator deforms accordingly. This requires the mirror to be much more flexible than the base; thus, the base should be made of very hard metal materials. The design difficulty of the deformable mirror mainly lies in selecting the actuator materials.

The actuator materials for separating the actuator continuous-surface deformable mirrors are generally piezoelectric materials (Pb (Zr, Ti,) O_3) or electrostrictive materials (Pb (Mg, Nb,) O_3). The actuators are made of these two materials; the former is called PZT, and the latter is called PMN [6]. When a driving voltage is applied to a deformable mirror, the displacement of the voltage-driven actuator changes, resulting in a deformation of the surface shape of the deformable mirror, which is similar to a Gaussian function [37]. In addition, the surface shapes of

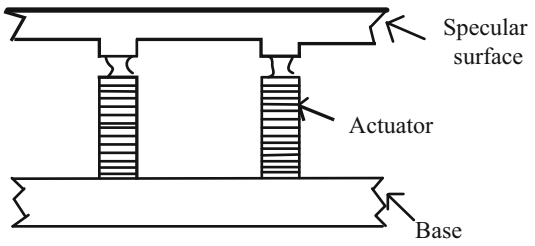


Fig. 27.8 Structure of continuous-surface deformable mirror for a separation actuator

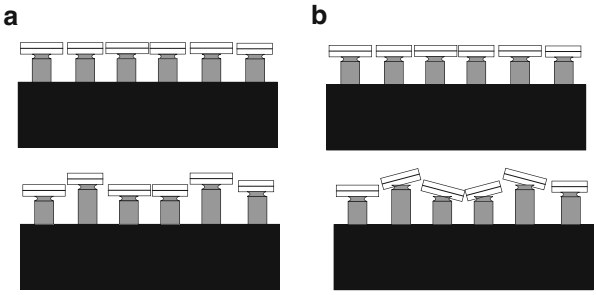


Fig. 27.9 Spliced deformable mirror. (**a**) Piston only; (**b**) piston and tip/tilt

mirrors on adjacent actuators also deform, and the deformation is generally 10–30% of the maximum deformation of deformable mirrors.

Spliced Deformable Mirror

The basic structure of the spliced deformable mirror is shown in Fig. 27.9. It can be observed from the figure that the mirror surface is separated into several small submirrors. One or three actuators are arranged under each submirror to adjust the mirror surface. If only one actuator is under the subreflector, the subreflector will move a piston along the beam-propagation direction to adjust the mirror surface. If three actuators are under the submirror, the submirror not only moves a piston along the beam-propagation direction, but also makes a 2D tip/tilt adjustment to change the surface shape of the deformable mirror.

A gap exists between each submirror of the deformable mirror, which reduces the utilization rate of the light energy received by the deformable-mirror surface and increases the difficulty of adjusting the mirror surface. To ensure a continuous wavefront structure, the phases between adjacent submirrors must be the same; therefore, this deformable mirror is rarely used in practical system applications.

Thin-Film Deformable Mirror

The thin-film deformable mirror is supported by multiple actuators, and the mirror surface is a continuous silicon-nitride mirror. Its structure is shown in Fig. 27.10a, and its deformation principle is shown in Fig. 27.10b. The film is very flexible; thus, only a small force is needed to change its surface shape. Generally, the actuator is made of an electrostrictive material, and a fixed support is required around the film to

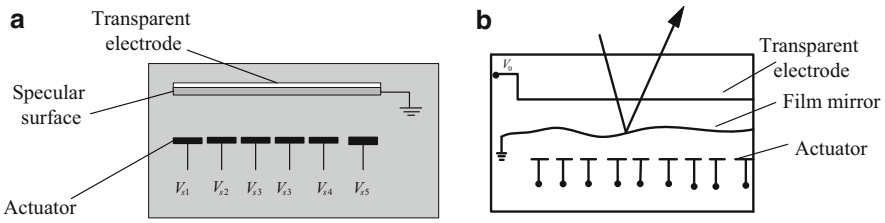


Fig. 27.10 (**a**) Structure diagram and (**b**) deformation schematic diagram of a thin-film deformable mirror

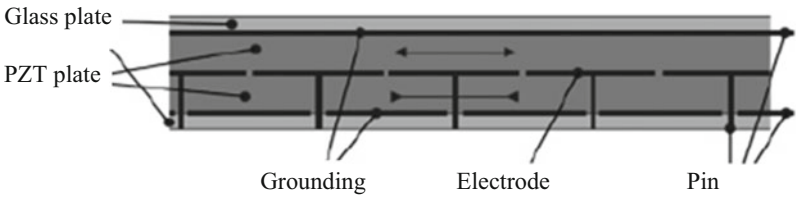


Fig. 27.11 Cross-section diagram of a dual-piezoelectric deformable mirror

provide tension for it to form a plane. The static deformation of the thin-film deformable mirror is [2]:

$$\nabla^2 z = \frac{-p(r)}{T_m}, \tag{27.1}$$

where $p(r)$ is the pressure on the film (unit: Nm^{-1}), and T_m is the linear tension of the film (unit: Nm^{-1}).

Dual-Piezoelectric Deformable Mirror

The dual-piezoelectric deformable mirror is composed of two piezoelectric-ceramic plates, two common electrodes (as shown in Fig. 27.11), one control electrode, and two thin glass plates. The cross-section of the deformable mirror is shown in Fig. 27.11. One common electrode is located above one piezoelectric-ceramic sheet, and the other is located below the other piezoelectric-ceramic sheet. The control electrode is located between the two piezoelectric-ceramic sheets. Two thin optical glass sheets are used as reflectors to connect the surfaces of the piezoelectric-ceramic sheets.

The basic principle of the deformable mirror is to apply a voltage to the control electrode using the transverse–converse piezoelectric effect of piezoelectric-ceramic materials. One of the piezoelectric-ceramic plates expands laterally, while the other shrinks laterally, resulting in a deformation of the thin lens connected to its surface. The deformation of a thin mirror can be expressed as [6]:

$$\nabla^4 r(x,y) = -A\nabla^2 V(x,y), \tag{27.2}$$

where $V(x, y)$ is the voltage on the piezoelectric-ceramic sheet, and A is a constant related to the sheet material properties.

MEMS Deformable Mirror

The micro-electromechanical system (MEMS) deformable mirror is a type of microdevice made of a silicon wafer using a thin-film process. The three manufacturing steps are depositing the film, patterning the film with a mask, and etching the film through the mask. These steps are repeated until a complete multilayer structure is generated. Finally, the silicon device is fabricated using wet etching. Figure 27.12 shows a discrete deformable mirror prepared using MEMS technology [38].

Voice-Coil Deformable Mirror

The bottom layer of the voice-coil deformable mirror, provided by the official website of the French ALPAO adaptive-optics product manufacturer, is composed of an independent coil, and the upper layer is composed of a covered thin mirror. The coil and thin mirror are connected using a magnet with a connecting rod. When the coil is energized, the magnet above the coil moves up and down under the reaction force of the ampere force, driving the connecting rod to move the thin mirror on top of it, causing the surface of the thin mirror to deform.

A voice-coil deformable mirror converts electrical energy (voltage) into kinetic energy (displacement). According to the static theory, the relationship between the voltage U and displacement x is [39]:

$$x = \frac{k_s U}{RK}, \tag{27.3}$$

where k_s denotes the motor-force constant, R is the coil resistance, and K is the structural stiffness matrix.

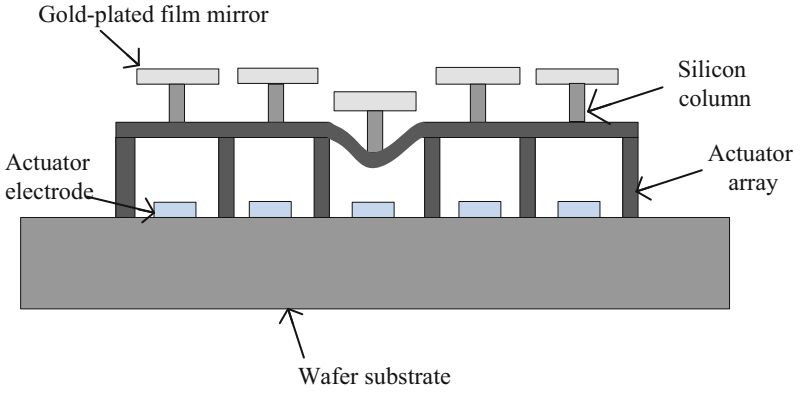


Fig. 27.12 MEMS discrete-surface deformable mirror

Performance Index of Deformable Mirrors

The main performance metrics of a deformable mirror, in addition to its luminous aperture and surface accuracy, are as follows:

1. Linearity: This mainly depends on the actuator material. The better the linearity of a deformable mirror, the higher its control accuracy.
2. Number of control units: This is the number of deformable-mirror actuators. It is usually determined by an adaptive-optical system according to the correction requirements and is proportional to the correctable spatial frequency.
3. Response time: When a control voltage is applied to the actuator of a deformable mirror, the time from the beginning to the end of the deformation of the surface shape is called the response time. The shorter the response time, the better the performance of the deformable mirror.
4. Resonant frequency: This is the lowest frequency before the deformable mirror reaches the resonant point. It is generally required to be ten times higher than the operating frequency of the deformable mirror.
5. Stability: The actuator of a deformable mirror must have sufficient strength, as well as static and dynamic stability. Taking discrete deformable mirrors as an example, static stability means that the deformable mirrors can maintain the same surface shape when working in different environments, whereas dynamic stability means that the surface can be restored to its original shape when the deformable mirrors stop working [40].

Model of a Deformable Mirror

We consider the relationship between the shape function $\varphi(x, y)$ of the deformable mirror, influence function $f_k(x, y)$ of the kth actuator, and control-voltage signal V_k of the kth actuator:

$$\varphi(x,y) = \sum_{k=1}^{n} V_k \cdot f_k(x,y), \tag{27.4}$$

where n represents the number of actuators. f is the influence function of the deformable mirror, which is usually regarded as a Gaussian function [37]:

$$f(x,y) = \exp\left[\mathrm{In}\omega\left(\sqrt{(x-x_i)^2 + (y-y_i)^2}/d\right)^{\alpha}\right], \tag{27.5}$$

where α is the exponent of the Gaussian function; d is the distance between actuators; x_i is the abscissa of the ith actuator; and y_i is the ordinate of the ith actuator; that is, (x_i, y_i) is the coordinate of the ith actuator. N is the number of actuator control units and ω is the coupling coefficient.

Deformable-Mirror Control Algorithms

A deformable mirror is one of many wavefront correctors. It can compensate for wavefront aberrations of a light wave by changing its optical path and correcting the wavefront aberration. We not only improved the structure and manufacturing process of deformable mirrors through technical innovation but also improved the correction accuracy of the deformable mirror by optimizing its control algorithm.

This section introduces a control algorithm, decoupling control algorithm, and sparse-sampling control algorithm based on the Prandtl–Ishlinskii (PI) hysteresis model and summarizes the application of deformable-mirror control algorithms in adaptive optics studied by Xi'an University of Technology.

Control Algorithm Based on the PI Hysteresis Model

The actuator of the piezoelectric deformable mirror, driven by the control voltage, cannot keep up with the change in the applied voltage and has hysteresis characteristics. Yang [41] proposed a control algorithm to eliminate the hysteresis characteristics and obtain the actual voltage, after the actuator eliminated the hysteresis. The basic principle of the algorithm is to describe the hysteresis characteristics of the piezoelectric deformable mirror using a mathematical model of the PI hysteresis, and then use its inverse model to eliminate the hysteresis characteristics of the piezoelectric deformable mirror.

The surface-shape information of a piezoelectric deformable mirror is expressed using a Zernike polynomial, as follows:

$$\varphi(x,y) = \sum_{j=1}^{m} a_j Z_j(x,y), \tag{27.6}$$

where a_j is the coefficient of the jth term of the Zernike polynomial and m is the number of terms in the Zernike polynomial. The influence function $f_i(x, y)$ of the ith actuator is expressed as a Zernike polynomial, as follows:

$$f_i(x,y) = \sum_{k=1}^{M} b_{ki} Z_k(x,y), \tag{27.7}$$

where b_{ki} denotes the coefficient of the ith actuator at $Z_k(x, y)$. Equation (27.6) can also be expressed as

$$\begin{aligned} \varphi(x,y) &= \sum_{i=1}^{N} v_i f_i(x,y) \\ &= \sum_{i=1}^{N} v_i \left[\sum_{k=1}^{M} b_{ki} Z_k(x,y) \right] \\ &= \sum_{k=1}^{M} \left(\sum_{i=1}^{N} v_i b_{ki} \right) Z_k(x,y). \end{aligned} \tag{27.8}$$

The Zernike polynomial coefficients of the influence functions of n piezoelectric actuators are stored in matrix B, which is called the response matrix. The relationship between the piezoelectric-actuator control-voltage vector $\boldsymbol{v}$ and deformable-mirror surface-shape vector $\boldsymbol{a}$ is obtained as follows:

$$\boldsymbol{a} = \boldsymbol{B}\boldsymbol{v}, \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \tag{27.9}$$

where n is the number of actuators and v_n is the driving voltage of the nth actuator. The deformation of each actuator is calculated using this formula. The mathematical model of a PI hysteresis is expressed as [41]

$$\begin{aligned} y(t) &= \boldsymbol{w}^{\mathrm{T}} \boldsymbol{H}_r[v(t), \boldsymbol{y}_0] \\ &= \sum_{i=0}^{n} w_i \cdot \max\{v(t) - r_i, \min[v(t) + r_i, y_i(t - T)]\}, \end{aligned} \tag{27.10}$$

where $w^T = [w_0, w_1, \ldots, w_n]$ is the weight vector, $\boldsymbol{H}_r$ represents a hysteresis operator with threshold r, $r = [r_0, r_1, \ldots, r_n]^T$ is the threshold vector, and $0 = r_0 < r_1 < \ldots < r_n < r_{\max}$ is the maximum input value. $n + 1$ is the number of operators used, v is the input (voltage) of the hysteresis operator, y_0 is the initial value of the output (deformation) of the hysteresis operator, T is the sampling period, and $y_i(t - T)$ is the output of the ith operator the moment before.

Because the PI hysteresis model and its inverse model are symmetric with respect to $y = x$, and their product is the unit matrix E, the expression of the output drive voltage (also the actual voltage after eliminating the hysteresis), with the deformation as the input of the PI-hysteresis inverse model, is

$$\begin{aligned} v(t) &= w'^{T} \boldsymbol{H}_{r'}[y(t), v_0] \\ &= \sum_{i=0}^{n} w'_i \cdot \max\{y(t) - r'_i, \min[y(t) + r'_i, v_i(t - T)]\}, \end{aligned} \tag{27.11}$$

where $w'^T = \left[w'_0, w'_1, \ldots, w'_n\right]$ is the weight vector of the inverse model, $r' = \left[r'_0, r'_1, \ldots, r'_n\right]^T$ is the threshold vector of the inverse model, $\boldsymbol{H}_{r'}$ represents a hysteresis operator with a threshold of r', and $v_i(t - T)$ is the output of the previous moment of the ith operator.

Decoupling Control Algorithm

When an actuator of a deformable mirror operates, the mirror connected to the actuator position deforms, and the mirrors connected to the adjacent actuators also deform to some extent, which causes a coupling between the actuators. When multiple deformable mirrors work synchronously, owing to the different spatial-

distribution characteristics of the deformable mirrors, a coupling effect occurs between the deformable mirrors. This may cause their correction quantities to offset each other and waste their correction range.

Coupling affects the correction effect of the deformable mirror and the system stability. To suppress this coupling, many scholars have proposed decoupling control algorithms to improve the correction accuracy of deformable mirrors.

Decoupling Control Algorithm of a Single Deformable Mirror

Zernike polynomials can only describe slowly changing wavefronts. Because the influence function of the deformable-mirror actuator changes rapidly with space in the local position, the influence function of the actuator cannot be expressed using Zernike polynomials. The accuracy of the deformable mirror is limited when determining the control quantity or driving voltage.

Because Zernike polynomials have difficulty expressing wavefront information completely and accurately, Liu et al. [42] proposed to directly use the wavefront-aberration plane to obtain the driving voltage of the deformable mirror. That is, they decouple the wavefront from the driving voltage through a direct-wavefront response matrix.

The surface shape of a specified area is expressed by Eqs. (27.4) and (27.5) as [42]

$$S = \sum_{i=1}^{n} v_i \times R_i, \tag{27.12}$$

where

$$S = \begin{bmatrix} s_{11}\ s_{12} \ldots s_{1e} \\ s_{21}\ s_{22} \ldots s_{2e} \\ \ldots\ \ \ldots\ \ \ldots\ldots \\ s_{h1}\ s_{h2} \ldots s_{he} \end{bmatrix}; R_i = \begin{bmatrix} r_{i_{11}}\ r_{i_{12}} \ldots r_{i_{1e}} \\ r_{i_{21}}\ r_{i_{22}} \ldots r_{i_{2e}} \\ \ldots\ \ \ldots\ \ \ldots\ldots \\ r_{i_{h1}}\ r_{i_{h2}} \ldots r_{i_{he}} \end{bmatrix},$$

where i represents the ith actuator, S represents the matrix of the specified area, and s_{11}, s_{12}, …, s_{he} are elements in matrix S, that is, the value of each element corresponding to the sampling point. h is the number of horizontal samples in the specified area, e is the number of vertical samples in the specified area, and the total number of samples is $h \times e$. R_i is also a matrix, where $r_{i_{11}}$, $r_{i_{12}}$, …, $r_{i_{he}}$ are the elements in matrix R_i, and each element corresponds to the value of the sampling point of the ith actuator in the specified area under the unit voltage.

The surface shape in the specified area can also be expressed by a Zernike expression

$$\Phi = \sum_{j=1}^{m} a_j \times Z_j, \tag{27.13}$$

where Φ is a matrix composed of the sampling points of the wavefront shape in the specified area, Z_j is a matrix composed of the values of Zernike polynomials in the specified size area, and a_j is the coefficient of the jth term of the Zernike polynomials.

According to the wavefront-compensation principle of deformable mirrors,

$$\Phi + S = 0, \tag{27.14}$$

where S is a zero matrix with the same dimensions. From Eqs. (27.12) and (27.14), we obtain

$$\Phi = -\sum_{i=1}^{m} v_i \times R_i. \tag{27.15}$$

Because the wavefront information can also be expressed by Zernike polynomials,

$$\sum_{j=1}^{m} a_j \times Z_j = -\sum_{i=1}^{m} v_i \times R_i. \tag{27.16}$$

It can be seen from Eq. (27.16) that no matter how many Zernike terms there are, the driving voltage of the deformable mirror can be calculated.

Decoupling Control Algorithm of Dual-Deformable Mirrors

Dual-deformable mirrors are used in the system to correct the wavefront distortion and improve the correction accuracy. The first type is a deformable mirror with a high coupling coefficient and low spatial resolution, commonly called a low-order deformable mirror, which is used to compensate for large-scale and low-order aberrations in wavefront distortion. The second type is a deformable mirror with a low coupling coefficient and high spatial resolution, usually called a high-order deformable mirror, which is used to compensate for small and high-order aberrations in wavefront distortion.

The residual-wavefront slope g is reconstructed using a Zernike polynomial, and the expression is [43]

$$g = Za, \tag{27.17}$$

where Z is the wavefront reconstruction matrix composed of Zernike polynomials and a is the coefficient of the Zernike polynomials. According to Eq. (27.19), the coefficient expression of the Zernike polynomial is

$$a = Z^{+}g \tag{27.18}$$

$$g = g_w + g_t, \tag{27.19}$$

where Z^+ is the generalized inverse matrix of Z, g_w is the uncompensated wavefront aberration of the lower-order deformable mirror, and g_t is the uncompensated wavefront aberration of the higher-order deformable mirror.

According to the space-compensation capability of the low-order deformable mirror, some low-order mode coefficients in coefficient a of the Zernike polynomial are selected as low-order aberrations and assigned to the low-order deformable mirror. The expression is

$$a_w = I_w a = I_w Z^+ g, \tag{27.20}$$

where I_w is the diagonal matrix used to select low-order aberrations. They are usually assigned to uncompensated mode aberrations of low-order deformable mirrors by setting the diagonal elements to 0 or 1, such as $I_w = \mathrm{diag}(1, 1, 1, 1, 1, 0, 0\ldots)$.

Because the low-order deformable mirror is responsible for compensating the low-order Zernike mode component in the wavefront aberration, the cumulative voltage v_w of the low-order deformable mirror is

$$v_w = R_w^+ Z a_w = R_w^+ Z I_w Z^+ g, \tag{27.21}$$

where R_w^+ is the pseudo-inverse matrix of the slope response of a low-order deformable mirror.

Using a typical proportional–integral (PI) controller, the driving voltage V_w of a low-order deformable mirror is

$$V_w(k+1) = k_p{}^* V_w(k) + k_i{}^* v_w, \tag{27.22}$$

where k_p and k_i are the control parameters of the PI controller, representing the proportional and integral coefficients of the current controller, respectively. Moreover, the wavefront-distortion slope g_t of a high-order deformable mirror is expressed as

$$g_t = g - g_w = (I - Z I_w Z^+) g, \tag{27.23}$$

where I denotes the identity matrix. The cumulative voltage v_t and driving voltage V_t of the high-order deformable mirror are as follows:

$$v_t = R_t^+ g_t = R_t^+ (I - Z I_w Z^+) g \tag{27.24}$$

$$V_t(k+1) = k_p{}^* V_t(k) + k_i{}^* v_t, \tag{27.25}$$

where R_t is the slope-response matrix of the high-order deformable mirror.

Because deformable mirrors have fitting errors for Zernike modes of different orders, coupling will occur when high-order and low-order deformable mirrors work simultaneously. To suppress the coupling between the two deformable mirrors, according to the correlation between the influence function of each actuator of the higher-order deformable mirror and the Zernike mode assigned to the lower-order

deformable mirror, a coupling-suppression matrix R_m of the higher-order deformable mirror was constructed. The suppression matrix is expressed as follows:

$$R_m(i,j) = k\frac{\iint V_i(x,y)Z_j(x,y)ds}{\iint Z_j(x,y)Z_j(x,y)ds}, \tag{27.26}$$

where V_i (x, y) is the influence function of the ith actuator of the high-order deformable mirror, $Z_j(x, y)$ is the surface function of the jth-order Zernike mode, and k is the coupling-suppression factor used to adjust the coupling-suppression degree of the high-order deformable mirror.

If the high-order deformable mirror is not coupled to the low-order deformable mirror, the total drive voltage satisfies the following formula:

$$0 = R_m V_t. \tag{27.27}$$

To eliminate the coupling component between the high-order and low-order deformable mirrors, we reset the driving voltage V_t of the high-order deformable mirror according to Eq. (27.27). The process is as shown in Eqs. (27.28) and (27.29):

$$\begin{bmatrix} I \\ R_m \end{bmatrix} V_t' \approx \begin{bmatrix} V_t \\ 0 \end{bmatrix} \tag{27.28}$$

$$V_t' = \begin{bmatrix} I \\ R_m \end{bmatrix}^+ \begin{bmatrix} V_t \\ 0 \end{bmatrix} = C \begin{bmatrix} V_t \\ 0 \end{bmatrix}. \tag{27.29}$$

In Eq. (27.29), C is the reset matrix and V_t' is the control voltage of the high-order deformable mirror; that is, there is no voltage component coupled with the low-order deformable mirror in V_t'.

Laplace-Eigenfunction Decoupling Control Algorithm

Because the first-order partial derivative of the Zernike polynomial is not orthogonal, the influence function of the higher-order deformable-mirror actuator must be accurately measured or fitted before constructing its coupling-suppression matrix. Some large-aperture telescope systems require numerous deformable mirrors with actuating elements. The large number of actuating elements makes it difficult to measure the influence function of each actuator. Therefore, a Laplace-eigenfunction decoupling control algorithm is introduced.

By solving the Laplace eigenfunctions in different pupil regions under different homogeneous Neumann boundary conditions, a wavefront-reconstruction matrix is constructed using the eigenfunctions, and the driving voltage of the higher-order deformable mirror is calculated. Because the first-order partial derivative of the eigenfunction is orthogonal, it is not necessary to measure the influence function of the higher-order deformable-mirror actuator when constructing the coupling-suppression matrix of the higher-order deformable mirror [44].

Under the homogeneous Neumann boundary condition, the first-order partial derivative of the Laplace eigenfunction $\nabla W(r)$ is orthogonal [43]:

$$\begin{aligned}\int_{\sigma} \nabla W_i(r) \nabla W_j(r) d\sigma &= -\int_{\sigma} \nabla W_j(r) \nabla^2 W_i(r) d\sigma \\ &= \gamma_i^2 \int_{\sigma} W_j(r) W_i(r) d\sigma \\ &= \gamma_i^2 c_i \delta_{ij} \\ &= \begin{cases} \gamma_i^2 c_i & i = j \\ 0 & i \neq j \end{cases},\end{aligned} \tag{27.30}$$

where c is the pupil boundary, σ is the pupil area, $W(r)$ is the eigenfunction, $\nabla W(r)$ is the first-order partial derivative of the eigenfunction, and γ^2 is the eigenvalue corresponding to the eigenfunction $W(r)$.

The eigenfunction is used to construct the wavefront-reconstruction matrix E instead of the Zernike wavefront-reconstruction matrix Z in section "Decoupling Control Algorithm of Dual-Deformable Mirrors." Taking the multistage corrector as an example, E can be expressed as follows:

$$E = \begin{Bmatrix} \frac{1}{A_1}\int_{P_1} \frac{\partial W_1(r)}{\partial x} d\sigma & \frac{1}{A_1}\int_{P_1} \frac{\partial W_2(r)}{\partial x} d\sigma & \cdots & \frac{1}{A_1}\int_{P_1} \frac{\partial W_{m_e}(r)}{\partial x} d\sigma \\ \frac{1}{A_2}\int_{P_2} \frac{\partial W_1(r)}{\partial x} d\sigma & \frac{1}{A_2}\int_{P_2} \frac{\partial W_2(r)}{\partial x} d\sigma & \cdots & \frac{1}{A_2}\int_{P_2} \frac{\partial W_{m_e}(r)}{\partial x} d\sigma \\ \vdots & \vdots & & \vdots \\ \frac{1}{A_N}\int_{P_N} \frac{\partial W_1(r)}{\partial x} d\sigma & \frac{1}{A_N}\int_{P_N} \frac{\partial W_1(r)}{\partial x} d\sigma & \cdots & \frac{1}{A_N}\int_{P_N} \frac{\partial W_{m_e}(r)}{\partial x} d\sigma \\ \frac{1}{A_1}\int_{P_1} \frac{\partial W_1(r)}{\partial y} d\sigma & \frac{1}{A_1}\int_{P_1} \frac{\partial W_2(r)}{\partial y} d\sigma & \cdots & \frac{1}{A_1}\int_{P_1} \frac{\partial W_{m_e}(r)}{\partial y} d\sigma \\ \frac{1}{A_2}\int_{P_2} \frac{\partial W_1(r)}{\partial y} d\sigma & \frac{1}{A_2}\int_{P_2} \frac{\partial W_2(r)}{\partial y} d\sigma & \cdots & \frac{1}{A_2}\int_{P_2} \frac{\partial W_{m_e}(r)}{\partial y} d\sigma \\ \vdots & \vdots & & \vdots \\ \frac{1}{A_N}\int_{P_N} \frac{\partial W_1(r)}{\partial y} d\sigma & \frac{1}{A_N}\int_{P_N} \frac{\partial W_1(r)}{\partial y} d\sigma & \cdots & \frac{1}{A_N}\int_{P_N} \frac{\partial W_{m_e}(r)}{\partial y} d\sigma \end{Bmatrix},$$

where A_k is the area of the kth effective aperture of the Hartmann sensor, p_k is the kth effective aperture of the Hartmann sensor, $k = 1, 2, \ldots, N$, and $\frac{\partial W_i(r)}{\partial x}$ and $\frac{\partial W_i(r)}{\partial y}$ represent the first-order partial derivatives of the ith eigenfunction in the x and y directions, respectively. m_e is the order of the eigenfunction and N is the number of effective apertures in the Hartmann sensor.

It can be seen from Eq. (27.30) that the first-order partial derivative of the eigenfunction is orthogonal. Eq. (27.31) is written in discrete form, as follows:

$$\sum_{k=1}^{N}\left(\frac{\partial W_i(x,y)_k}{\partial x}\frac{\partial W_j(x,y)_k}{\partial x}+\frac{\partial W_i(x,y)_k}{\partial y}\frac{\partial W_j(x,y)_k}{\partial y}\right)=0, \quad i\neq j, \tag{27.31}$$

where $\frac{\partial W_i(x,y)_k}{\partial x}$ represents the first partial derivative of the eigenfunction in the x direction at the kth sampling point, also known as the slope, and $\frac{\partial W_i(x,y)_k}{\partial y}$ represents the first partial derivative of the eigenfunction at the kth sampling point in the y direction, namely the slope.

For Eq. (27.31), when the sampling rate of the Hartmann sensor is high, the average slope on the subaperture is closer to the slope at the discrete sampling points, namely

$$\lim_{A_k\to 0}\frac{1}{A_k}\int_{p_k}\frac{\partial W_i(r)}{\partial x}d\sigma=\frac{\partial W_i(x,y)_k}{\partial x} \tag{27.32}$$

$$\lim_{A_k\to 0}\frac{1}{A_k}\int_{p_k}\frac{\partial W_i(r)}{\partial y}d\sigma=\frac{\partial W_i(x,y)_k}{\partial y}. \tag{27.33}$$

Equations (27.32) and (27.33) indicate that the column vector of the wavefront-reconstruction matrix E corresponding to the eigenfunction is orthogonal when the sampling frequency of the Hartmann sensor subaperture is high. Therefore, Eq. (27.26) can also be written as

$$R_m(i,j)=k\frac{\sum_{r=1}^{2N}R_t(r,i)E(r,j)}{\sum_{r=1}^{2N}E(r,j)E(r,j)}, \tag{27.34}$$

where $R_t(r, i)$ represents the element corresponding to position (r, i) in the slope-response matrix of the high-order deformable mirror, and $E(r, j)$ represents the element corresponding to position (r, j) in the eigenfunction.

The new matrix R_m is obtained by substituting Eq. (27.34) into Eqs. (27.28) and (27.29), and the reset matrix of the higher-order deformable mirrors based on the Laplace-eigenmode driving voltage is obtained; that is, the coupling component between the higher-order and lower-order deformable mirrors is eliminated.

Sparse-Sampling Control Algorithm

To accurately control the high-order and low-order coefficients of deformable-mirror shape expressions, Cheng et al. [45] proposed a sparse-sampling control algorithm to control a deformable mirror.

A deformable mirror typically consists of multiple actuators and a reflector. When the actuator applies the driving voltage, the surface shape of the deformable mirror changes accordingly. The driving voltage of each actuator is arranged to form the control-parameter vector S, which is [45]

$$S = [s_1\ s_2 \ldots s_i \ldots s_n]^T, \tag{27.35}$$

where n is the number of actuators, i is the number of the ith actuator (i.e., the actuator), s_i is the driving voltage of the ith actuator, and $s_i \in [-1, 1]$.

After the control-parameter vector S is applied, the surface shape of the deformable mirror is represented by the sum of Zernike polynomials, as follows:

$$z(x, y) = \sum_{1}^{K=m} c_K Z_K(x, y). \tag{27.36}$$

The formula is a Zernike polynomial and c_k is the coefficient of the polynomial. The value of m is related to the fitting accuracy and the solution complexity of the surface shape. The relationship between the two should be comprehensively considered and an appropriate value should be selected.

Because Eq. (27.36) is a definite expression, its coefficient component vector is used to describe the surface-shape information of the deformable mirror. The coefficient vector of the Zernike polynomial is defined as the surface-shape feature vector c, which is expressed as

$$c = [c_1\ c_2 \ldots c_K \ldots c_m]^T. \tag{27.37}$$

Assuming that the deformation of the deformable mirror is within the linear-variation range, and considering the linear mechanical model of the deformable mirror, the relationship between the control-parameter vector of the deformable mirror and the surface-shape feature vector is as follows:

$$c = Gs, \tag{27.38}$$

where G is the response matrix of the deformable mirror.

In the iterative control process of the deformable mirror, the feature vector of the current surface shape is represented by c_N and the feature vector of the target surface shape is represented by c_d, where N is the number of iterations. The mirror-shape correction is expressed as follows:

$$\Delta c_N = c_d - c_N. \tag{27.39}$$

From Eqs. (27.35), (27.38), and (27.39), it can be seen that the relationship between the change of the deformable-mirror surface shape and the change of the control parameters is as follows:

$$\Delta c_N = G\Delta s_N. \tag{27.40}$$

Because the response matrix G of the deformable mirror is irreversible and the number of columns is greater than the number of rows, the solution of Eq. (27.41) is essentially the solution of the underdetermined equations. Therefore, the sparse-control parameter φ can be obtained by establishing the sparse matrix θ_N obtained from the discrete cosine transform. The control-parameter variable Δs_N is sparsely sampled using the orthogonal matching-pursuit method, and is expressed as follows:

$$\Delta c_N = G\varphi\theta_N. \tag{27.41}$$

After completing the sparse solution of the control parameters, the sparse matrix of Eq. (27.41) is used to recover the changes of the actual control parameters φ, and the correction of the control parameter Δs_N is obtained. The expression is

$$\Delta s_N = \varphi\theta_N. \tag{27.42}$$

In summary, the iterative relationship between the control parameters of the deformable mirror is

$$s_{N+1} = s_N + \varphi\theta_N. \tag{27.43}$$

The mapping relationship of Eq. (27.38) can be written as

$$\Delta c = G\Delta s, \tag{27.44}$$

where Δs is the vector that controls the change in the parameters and Δc is the vector that changes the shape of the deformable mirror.

The zero-state surface shape c_{-1} is defined as the average surface shape of the deformable mirror before loading the control parameter; the parameter is defined as the zero-state control parameter s_{-1}. To reduce the repeated influence of hysteresis on the surface-shape data and improve the surface-shape accuracy, the following methods were used to obtain sample data.

At the beginning of the experiment, the deformable mirror was loaded with control parameter s_{-1}. Second, the control-parameter vector s_t, satisfying a random Gaussian distribution, was loaded, where t is the serial number of the surface-shape information collection. Using an interferometer, the surface-shape feature vector c_t of more than 70% of the effective area of the deformable mirror was measured, and the change of the control parameter Δs_t and the change of the surface-shape feature vector Δc_t were obtained. The expressions are

$$\Delta c_t = c_t - c_{-1} \qquad \Delta s_t = s_t - c_{-1}. \tag{27.45}$$

The sample data were collected, and the control parameter-change sample set $\Delta s'$ and surface-shape feature-vector change sample set $\Delta c'$ were established. The expressions are

$$\begin{aligned} \Delta s' &= [\Delta s_1\ \Delta s_2 \ldots \Delta s_R] \\ \Delta c' &= [\Delta c_1\ \Delta c_2 \ldots \Delta c_R], \end{aligned} \tag{27.46}$$

where R is the upper limit of t.

Using the sample set, the expression is

$$\min\left\{ \left(\|\Delta c' - G\Delta s'\|_F \right)^2 \right\}. \tag{27.47}$$

where $\|.\|_F$ is the F norm. The least squares method was used to solve Eq. (27.47), and the response matrix G was obtained as

$$G = \Delta c \left[\Delta s^T \left(\Delta s \Delta s^T \right)^{-1} \right]. \tag{27.48}$$

After Eq. (27.48) was substituted into Eq. (27.44), both sides of the equation were multiplied by the pseudo-inverse matrix G^+ of the response matrix, and the correction value of the control parameters was obtained as follows:

$$\Delta s_N = \left[\Delta c \left(\Delta s^T \left(\Delta s \Delta s^T \right)^{-1} \right) \right]^+ \Delta c_N. \tag{27.49}$$

During the experiment, the optimization objective in the experiment process was generally set to minimize the relative change of the deformable-mirror surface shape, and the initial value of the iteration was set as the median of the control parameters. At this time, the deformable mirror was in the middle state, under the given control parameters, and the eigenvector of the surface shape was defined as the initial value eigenvector c_0. The two groups of control parameters used in the experiment were $s_i \in [-1, 1]$; thus, the initial value of the iteration was the zero vector.

Application of Deformable-Mirror Control Algorithms in Adaptive-Optics Systems

In recent years, Ke Xizheng et al. at Xi'an University of Technology have devoted themselves to studying deformable-mirror control algorithms and applying them to adaptive-optics systems.

In 2018, Wu [46] investigated the stochastic parallel gradient-descent (SPGD) algorithm, studied wavefront-distortion correction using this algorithm, and analyzed the influence of different parameter values in the SPGD algorithm on the correction effect, and numerically simulated different parameter values.

The simulation results showed that when the disturbance amplitude $\delta u = 0.006$, the SPGD algorithm based on the variable-gain coefficient effectively corrected the

wavefront distortion. The SPGD algorithm also well corrected the wavefront distortion under different turbulence intensities. In addition to verifying the algorithm through numerical simulations, experiments were also conducted to correct the wavefront distortion at 600 m, 1300 m, and 10.2 km in an external field. This verified that the algorithm had a good wavefront-distortion correction ability.

In 2019, Han [47] used the SPGD algorithm to control a deformable mirror to correct optical-wavefront distortions. The simulation results showed that the performance-evaluation function's Strehl ratio improved from 0.15–0.87, and the beam distortion was improved. A deformable mirror was used as the wavefront corrector of the system and a wavefront-free adaptive system was built for the experiment. The experimental results showed that the deformable mirror had a good correction effect, and after the correction, the shape and brightness of the light spot were improved.

In 2020, Zhang [35] studied the proportional–integral–differential (PID) algorithm, and then analyzed and simulated the influence of the input universe of different control voltages and the input universe of the first derivative of different control voltages on the algorithm's correction results. The simulation results showed that the input universe of the control voltage should be reasonably selected, according to the actual wavefront situation, and that the actual input universe should be smaller than the set input universe to ensure the correction accuracy of the deformable mirror.

To verify the influence of the algorithm parameters on the correction results, an outfield experiment was conducted. The experimental results were consistent with the simulation results, proving the correctness of the dual fuzzy PID algorithm.

In 2020, Li [48] deduced the eigenmode of a deformable mirror, studied the correction principle of the eigenmode deformable-mirror method, established a mathematical model of the 69-element continuous-surface deformable mirror, and deduced its eigenmode. The wavefront distortion was corrected, based on the eigenmode method, and a numerical simulation was performed.

The simulation results showed that the stability of the system could be improved by improving the atmospheric environment. To study the performance of the model in a practical application system, wavefront distortions under different turbulence intensities were corrected, and an indoor experimental system was established.

The results showed that the model method could effectively correct wavefront distortions caused by atmospheric turbulence. An outdoor experimental system was established to conduct the experiment. The experimental results showed that the deformable-mirror eigenmode method effectively corrected the wavefront distortion caused by external-field dynamic turbulence.

In 2021, Cheng [49] studied a residual-distortion correction algorithm based on a PID algorithm. This algorithm mainly eliminated the residual part of the distortion variables, except for period expansion, and was used to simulate wavefront correction with large distortions, which improved the system performance. To verify the algorithm, an experimental system was built to analyze and correct the algorithm for an indoor static distorted wavefront, and 1.2-km, 10-km, and 100-km dynamic distorted wavefronts in the field, proving the correctness of the algorithm.

Summary

The wavefront corrector is an important wavefront-correction component in adaptive-optics systems, and its performance directly determines the correction level of the system. Several typical deformable mirrors were introduced, applied, and verified in practice and achieved good results.

However, owing to the different characteristics and manufacturing processes of deformable mirrors, the most suitable systems should be comprehensively considered when designing systems. Moreover, the constant improvement of application requirements and the emergence of high-integration technology require continuous innovation of technology and theory, as well as gradual improvements and enhancements of deformable-mirror control algorithms to improve the correction accuracy and communication performance of adaptive-optical systems [6].

China has made great breakthroughs in the research fields of deformable mirrors and their control algorithms. The main research units are the Institute of Optoelectronics Technology, Chinese Academy of Sciences, Changchun Institute of Optics and Mechanics, Chinese Academy of Sciences, and Xi'an University of Technology. Over the years, significant progress has been made in the research on deformable-mirror systems and element technology. However, many aspects remain to be studied in the future.

1. New structures and new fabrication processes should be studied to find new ways to improve deformable-mirror performance.
2. Actuator materials should be improved and actuator arrays with smaller spacing and higher density should be developed to meet the needs of larger telescopes.
3. The coating range of MEMS deformable mirrors should be studied to obtain good deformable-mirror shape quality [8].
4. The quantitative relationship between the number of actuators and number of sampling points for the decoupling control algorithm of a single deformable mirror should be studied to improve the algorithm solution accuracy [42].
5. Control algorithms for deformable mirrors should be optimized, simpler and more efficient control algorithms should be sought, and the deformable-mirror correction level and system stability should be improved.
6. The working bandwidth of dual deformable-mirror adaptive-optical systems should be improved to obtain a better aberration-compensation effect [43].

References

1. Babcock H W. The possibility of compensating astronomical seeing [J]. Publ Astron Soc Pacific, 1953, 65(386): 229–236.
2. Yellin M. Using membrane mirrors in adaptive optics [J].Imaging through the atmosphere. SPIE, 1976, 75(22): 97–102.
3. Kokorowski S A. Analysis of adaptive optical elements made from piezolectric bimorphs [J]. JOSA, 1979, 69(1):181–187.

4. Schwartz C, Ribak E, Lipson S G. Bimorph adaptive mirrors and curvature sensing [J]. JOSA A, 1994, 11(2):895–902.
5. Feeman, R. H, Garcia, et al. High-speed deformable mirror system [J]. Applied Optics, 1982, 21(4):589–595.
6. Lin Xudong, Xue Chen, Liu Xinyue, et al. Development Status of Adaptive Optics Wavefront Corrector Technology [J]. China Optics, 2012, 5 (04): 337–351.
7. Hulburd W G, Barrett T K, Cuellar E L, et al. High-bandwidth long-stroke segmented mirror for atmospheric compensation [J].SPIE, 1992, 1543(38):64–75.
8. Takami H, Iye M. Membrane deformable mirror for SUBARU adaptive optics [J]. SPIE, 1994, 2201(40):762–767.
9. Oppenheimer B R, Palmer D L, Dekany R G, et al. Investigating a Xinetics Inc.deformable mirror [J]. SPIE, 1997, 3126(43):569–579.
10. Bifano T G, Krishnamoorthy Mali R, Dorton J K, et al. Continuous-membrane surface-micromachined silicon deformable mirror [J]. Optical Engineering, 1997, 36(5):1354–1360.
11. Lowrey W H, Wynia J L, Ealey M A, et al. Characterization of three advanced deformable mirrors [J]. SPIE, 1998, 3433(44):388–393.
12. Spinhirne J M, Ameer G A. Adaptive optics using the 3.5-m Starfire Optical Range telescope [J]. SPIE, 1997, 3126(43): 257–268.
13. Paterson C, Munro I, Dainty C. A low cost adaptive optics system using a membrane mirror [J]. Optics express, 2000, 6(9):195–185.
14. Fernández E J, Iglesias I, Artal P. Closed-loop adaptive optics in the human eye [J]. Opt.Lett, 2001, 26(10):746–748.
15. Hull T, Trauger J, Macenka S, et al. Examination of specific telescope constraints and optimum architectures for very high contrast eclipse imagery of extra-solar planets [J]. SPIE, 2002, 4860(48):49–57.
16. Ealey M A, Trauger J T, Macewen H A. High density deformable mirrors to enable coronagraphic planet detection [J]. SPIE, 2004, 5166(50):192–199.
17. Oya S, Bouvier A, Guyon O, et al. Performance of the deformable mirror for Subaru LGSAO [J]. SPIE, 2006, 6272(52):62724S1–62724S8.
18. Comelissen S A, Berden P A, Beano T G. A 4096 element continuous facesheet MEMS deformable mirror for high-contrast imaging [J]. SPIE, 2008, 6888(53):238–247.
19. Dekany R.Innovative deformable mirror for Palomar Observatory [J]. SPIE Newsroom, 2010, 10(2):1–3.
20. Verpoort S, Rausch P, Wittrock U. Characterization of a miniaturized unimorph deformable mirror for high powercw-solid state lasers [J]. SPIE, 2012, 8253(58):67–68.
21. Kuiper S, Doelman N, Nieuwkoop E, et al. Electromagnetic deformable mirror development at TNO [J]. SPIE, 2016, 9912(10): 24–30.
22. Kuiper S, Doelman N, Human J, et al. Advances of TNO's electromagnetic deformable mirror development [J]. SPIE, 2018, 10706(19): 345–352.
23. Jiang Wenhan, Wu Xubin, Ling Ning, et al. 37 unit adaptive optical system [J]. Optoelectronic Engineering, 1995, 22 (1): 38–45.
24. Xingzheng Ke,Ke Dong.Optical Wireless Communication[M], Beijing Springer Press,2022
25. Ling Ning, Guan Chunlin, Wang Lan, et al. Development of 61 unit discrete piezoelectric deformable mirror [J]. Intense Laser and Particle Beam, 1998, 10 (4): 48–51.
26. Xiang Dong, Chen Haiqing, Wang Qingling. Development of deformable mirror with transparent electrode [J]. Optoelectronic Laser, 2006, 16 (6): 678–683.
27. Zhou Hong, Ning Yu, Guan Chunlin, et al. Design and development of dual piezoelectric deformable mirror [J]. Journal of Optics, 2009, 29 (6): 1437–1442.
28. Wei K, Zhang X J, Xian H, et al. First light on the 127-element adaptive optical system for 1.8 m telescope [J]. Chinese Opt. Lett., 2010, 8(11):1019–1024.
29. Lin Xudong, Liu Xinyue, Wang Jianli, et al. Research on flattening calibration of deformable mirror based on interferometer measurement [J]. Acta Photonica Sinica, 2012, 15 (2): 193–199.
30. Lin Xudong, Liu Xinyue, Wang Jianli, et al. Development of continuous mirror deformable mirror based on piezoelectric ceramic actuator [J]. Progress in Laser and Optoelectronics, 2014, 51 (9): 31–40.

31. Lin Xudong, Liu Xinyue, Wang Jianli, et al. 137 Unit deformable mirror performance test and correction ability experiment [J]. Optical Precision Engineering, 2013, 21 (2): 267–273.
32. You Juncheng, Zhou Hong, Guan Chunlin, et al. Performance test and analysis of 913 deformable mirrors [J]. Laser Technology, 2019, 41 (6): 867–871.
33. Du Yimin, Gao Yan, Xu Boqian, etc Integrated design of spliced sub mirror module based on combination optimization algorithm [J]. Acta Photonica Sinica, 2022, 51 (4): 293–302.
34. Dong Yulei Research on real-time performance of wave front processing algorithm of adaptive optical system based on GPU [D] Changchun: University of Chinese Academy of Sciences (Changchun Institute of Optics, Precision Machinery and Physics, Chinese Academy of Sciences), 2018.
35. Zhang Danyu. Adaptive optics wavefront distortion control and experimental research [D]. Xi'an: Xi'an University of Technology, 2023.
36. Qiao Ruijie, Chen Mingsong Research on deformable mirror in adaptive optics [J]. Electronic Design Engineering, 2010, 18 (3): 74–75.
37. Guo Youming, Rao Changhui, Bao Hua. A direct calculation method for the response matrix of adaptive optical systems [J]. Journal of Physics, 2014, 63 (14): 455–461.
38. Bifano T. Adaptive imaging: MEMS deformable mirrors[J]. Nature Photonics , 2011, 5(1): 21–27.
39. Cao Shuo. Performance optimization and multi field coupling analysis of voice coil deformable mirror [D]. Wuxi, Jiangsu: Jiangnan University, 2024.
40. Ling Ning, Guan Chunlin. Development of deformable mirrors [J]. Optoelectronic Engineering, 1995, 22 (1): 14–26.
41. Yang Zongfeng. Research on control method of piezoelectric deformable mirror based on neural network and hysteresis model [D]. Ningbo, Zhejiang: Ningbo University, 2021.
42. Liu Cheng, Yu Fei, Ding Lin, et al. A new decoupling control method for continuous deformable mirror [J]. Aerospace Return and Remote Sensing, 2019, 40 (02): 89–98.
43. Cheng Tao Research and application of multi-level corrector decoupling control technology [D]. Beijing: University of Chinese Academy of Sciences (Institute of Optoelectronics Technology, Chinese Academy of Sciences), 2018.
44. Cheng Tao, Liu Wenjin, Yang Kangjian, etc. Decoupling control algorithm of Woofer Tweeter adaptive optical system based on Laplace eigenfunction [J]. China Laser, 2018, 45 (9): 295–304.
45. Cheng X M, Zhao Y J, Yan L, et al. Sparse sampling algorithm for deformable mirror control [J]. Journal of Optics, 2020, 22(3): 035704–035704+9.
46. Wu Jiali. Experimental study on coherent optical communication system without wavefront detection [D] Xi'an: Xi'an University of Technology, 2018.
47. Han Kona. Comparative study of deformable mirror and liquid crystal spatial light modulator for wavefront correction [D]. Xi'an: Xi'an University of Technology, 2021.
48. Li Mei. Experimental study on correction of beam wavefront distortion by eigenmode method [D]. Xi'an: Xi'an University of Technology, 2023.
49. Cheng Shuang. Research on correction technology of large amplitude distorted wavefront in free space coherent optical communication [D]. Xi'an: Xi'an University of Technology, 2024.

Liquid-Crystal Spatial Light Modulators and Their Applications

28

Contents

Introduction 1016
Liquid-Crystal Spatial Light Modulators 1016
Liquid Crystal and Spatial Light Modulators 1016
Research Progress 1018
LC-SLM Basic Working Principle 1021
Electronically Controlled Birefringence Properties of Liquid Crystals 1021
Directional Distribution of Liquid Crystals 1025
Reasons for and Experiments with the Phase Calibration of Liquid-Crystal Spatial Light Modulators 1026
Double-Slit Interference 1027
Twyman–Green Interference (Fig. 28.7) 1028
Turbulence-Simulation Principle of a Liquid-Crystal Spatial Light Modulator 1028
Wavefront-Distortion Correction Principle of a Liquid-Crystal Spatial Light Modulator 1030
Classification of Liquid-Crystal Spatial Light Modulators 1031
Research Status of Optical-Modulation Technology 1032
Foreign Research Progress 1032
Domestic Research Status 1036
Performance Evaluation of Liquid-Crystal Spatial Light Modulators 1038
Application of Liquid-Crystal Spatial Light Modulators to Adaptive Optics 1040
Phase-Screen Simulation of a Liquid-Crystal Spatial Light Modulator 1040
LC-SLM Coherent Optical Communication Wavefront-Correction Technology 1041
Summary 1058
References 1059

Abstract

Liquid-crystal spatial light modulators control the optical path of light waves by modulating the refractive index. They play an important role in adaptive optics as phase-correction devices. This chapter expounds on the domestic and foreign development and research status of liquid-crystal spatial light modulators and light-beam modulation. The principles of optical-amplitude and phase modulation of liquid-crystal spatial light modulators are introduced, and their future applications are proposed.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_28

Keywords

Adaptive optics · Liquid-crystal spatial light modulator · Wavefront distortion correction · Design of adaptive optical system

Introduction

Liquid-crystal spatial light modulators achieve control of the light path by modulation of the refractive index. As an important phase-correction device, it plays an important role in adaptive optics. This chapter describes the foreign development status of liquid-crystal spatial light modulator and the current status of foreign research on beam modulation of liquid-crystal spatial light modulator, introduces the realization principle of optical amplitude and phase modulation of liquid-crystal spatial light modulator, and looks forward to the future application of liquid-crystal spatial light modulator.

Liquid-Crystal Spatial Light Modulators

Liquid Crystal and Spatial Light Modulators

As condensed matter, the characteristics and structure of a liquid crystal are between those of solid crystals and isotropic liquids; it is an ordered fluid. In terms of macroscopic physical properties, it not only has the fluidity and viscosity of a liquid, but also the anisotropy of a crystal. Similar to crystals, it can produce birefringence, Bragg reflection, diffraction, and optical-rotation effects, and can also produce thermo-optical, electro-optical, or magneto-optical effects under the action of external fields.

An optical wavefront propagating through atmosphere will be distorted by atmospheric turbulence. An adaptive-optics system is an important method for suppressing the influence of atmospheric turbulence. The wavefront corrector is a key device in an adaptive-optics system, and its main function is to correct the distortion aberration measured by the wavefront sensor [1].

When light is incident on a liquid crystal, the liquid-crystal molecules have two refractive indices, which are parallel and perpendicular to the optical axis. This is called the birefringence characteristic. According to the anisotropy characteristics of liquid-crystal molecules, the refractive index of the light beam differs in all directions when the liquid-crystal molecules propagate.

When a light beam propagates in liquid-crystal molecules, two types of light waves are often formed: ordinary and extraordinary. The refractive indices of the two are n_o and n_e, respectively. The transmission rate differs between the two, resulting in significantly different refractive-index characteristics, which is called birefringence [2, 3].

A liquid-crystal spatial light modulator can modulate the wavefront phase; it has no mechanical inertia, extremely low power consumption, an extremely low energy-absorption rate, and a high system resolution. In its liquid-crystal molecule unit, the

voltage can be programmed to control the liquid-crystal molecules. Therefore, it has unique advantages for wavefront correction [4–6].

Spatial light modulators can adjust the light beam and load information in one- or two-dimensional optical data fields, and can effectively utilize the inherent speed, parallelism, and interconnection capabilities of light to modulate the light wave. Spatial light modulators contain many independent units, which are spatially arranged in one- or two-dimensional arrays. Each unit can be independently controlled by optical or electrical signals, using various physical effects, such as the Pockels, Kerr, acousto-optic, magneto-optic, semiconductor self-electro-optic, and photorefractive effects. The units change their own optical properties to modulate the light waves incident on them.

These individual small units are often referred to as the "pixels" of a spatial light modulator. The signal that controls the pixel is called the "write light." The input light wave that illuminates the entire device and is modulated is called the "read out light." The light wave emitted after passing through the spatial light modulator is called the "output light." Visually speaking, a spatial light modulator can be regarded as a transparent sheet whose transmittance or other optical-parameter distribution can be quickly adjusted as needed. The write signal contains information that controls the individual pixels of the modulator. The process of passing this information separately to the corresponding pixel positions is called "addressing."

A liquid-crystal spatial light modulator uses the liquid-crystal layer as the light-modulation material, which adopts the mixed field-effect working mode of a nematic liquid crystal. By applying different electric fields to each region of the crystal layer, the arrangement direction and position of the liquid-crystal molecules can be changed. This changes their optical properties, making them transparent or opaque, and modulates the optical signals.

Spatial light modulators are composed of cells controlled by many integrated circuits, and use electrical signals to control the 1 or 0 state of each cell. Liquid crystals are used as the modulation medium to form low-power, low-voltage spatial light modulator light-addressable liquid-crystal light valves. (They are typically called spatial light modulators when used for light operations and light valves when used in displays.) Electrically addressable thin-film transistors drive the liquid-crystal displays (TFT-LCDs).

The amount of phase correction by the liquid-crystal spatial light modulator is usually determined by the refractive index of the liquid-crystal material and the thickness of the liquid-crystal layer. In an actual operation, the liquid-crystal spatial light modulator corrects the front of the wave by changing the refractive index. The phase-modulation amount is approximately 1 mm, and this value is determined according to the thickness of the liquid-crystal layer constituting the modulator.

The lack of sufficient phase modulation limits the use of liquid-crystal spatial light modulators. By increasing the thickness of the liquid-crystal layer and improving the performance of the liquid-crystal material, the modulation amount with the refractive-index difference is improved. Although these changes increase the response time of the liquid-crystal spatial light modulator, the phase-modulation amount of the corrector can be greatly improved by adding an intrinsic phase-modulation amount of one wavelength to the liquid crystal through phasogram technology.

When an adaptive-optics system works in the visible-light band, the atmospheric coherence length is significantly reduced because of the shorter wavelength in the visible-light band. This places higher requirements on an adaptive-optics system composed of deformable mirrors.

Compared with a traditional deformable-mirror adaptive-optics system, the advantages of the liquid-crystal spatial light modulator, in correcting the wavefront in an adaptive-optics system, are as follows: The pixel density of the liquid-crystal spatial light modulator is high, and it is easy to achieve the order of millions of pixels, plus small pixels. Combined with the phase-hologram technology of binary optics, correction amounts on the order of more than 10 μ can be reached, which is better than the correction amount of deformable mirrors. In addition, the liquid-crystal spatial light modulator is stable, the device technology is excellent, and its cost performance is better than that of a deformable mirror.

Unfortunately, the liquid-crystal spatial light modulator has shortcomings in engineering applications, mainly reflected in its low energy utilization and slow response speed. The main reasons for the low energy-utilization rate are the polarization dependence of the liquid-crystal corrector and the dispersion effect of the liquid-crystal material. The liquid-crystal adaptive-optics system can only work on polarized light; thus, it can only correct the wavefront after it has been filtered by a polarizer, which results in energy loss. The dispersion effect of the liquid-crystal materials also restricts the energy-utilization rate.

The wavelength of the incident rays affects the refractive index of the material. This phenomenon makes it impossible to consistently correct distortions over a wide wavelength range for devices that change the refractive index to adjust the wavefront. In addition, the liquid-crystal corrector is a binary-diffraction correction device, and its diffraction-dispersion effect further restricts the working-band range of the system. These two factors reduce the spectral width of the system working band to approximately 10–30 nm and the optical-energy loss is serious.

The response speed of the liquid-crystal spatial light modulator also affects the working performance of the system. When dual-frequency liquid crystals are used to shorten the response time of the liquid-crystal spatial light modulator, the response time is on the order of sub-milliseconds. However, when driving a dual-frequency liquid crystal, a large external voltage must be connected, so this method is not widely used. Nematic liquid-crystal materials are used in liquid-crystal spatial light modulators. Reducing the thickness of the liquid-crystal layer shortens the response time, which reduces the impact of the response speed on the working performance to a certain extent.

Research Progress

Foreign Research Progress

In 1975, Meyer et al. discovered ferroelectric liquid-crystal materials with many novel electro-optical properties. Clark et al. proposed a surface-stable liquid-crystal mode and applied ferroelectric liquid crystals to the display field [7].

At the end of the twentieth century, the United States began a program code-named "Steered Agile Beams" to study liquid-crystal optical correctors. Major institutions in the United States also began to study liquid-crystal optical correctors [8, 9].

In 1989, Raytheon Corporation of the United States launched the first liquid-crystal optical phased array. In 1996, a reflective liquid-crystal phased array with 5000 independent electrodes and a $2 \times 2\text{-cm}^2$ device aperture size was launched. The Stokes–MacCulbrough measurement technique was used for measuring 10.6-μm beams of light. It was found that the device performance of the phased array was affected by the thermal effect, and $\pm 5^\circ$ field-of-view scanning was achieved [10].

In 2002, Rockwell Scientific Corporation of the United States developed a reflective liquid-crystal phased array with 512 independent electrodes. The array-element size was 10 μm and the optical aperture was $5 \times 5\ \text{mm}^2$. In 2005, a transmissive liquid-crystal optical phased array with a dual-frequency liquid-crystal material was proposed, with four independently addressable $1 \times 4\ \text{cm}^2$ areas. They were arranged in a linear array to produce a continuous $4 \times 4\ \text{cm}^2$ clear aperture with 8-μm pixel spacing. Experiments confirmed that under incident rays with a 1.55-μm wavelength, the field-of-view deflection angle was 2.5°, and the diffraction efficiency reached more than 60% [11].

In Reference [12], an optical phased array of $1 \times 12{,}288$ pixels was proposed. In this design, the pixel pitch was reduced from 1.8 μm to 1.6 μm. The backplane voltage increased from 5 V to 13 V and the effective light aperture increased from $7.4 \times 6.0\ \text{mm}^2$ to $21.66 \times 21.66\ \text{mm}^2$. Under the new liquid-crystal construction, when calibrated with the new automatic-correction program, the scanning angle of the optical phased array was more than doubled, and the response time was 5–30 ms. Operating wavelengths ranged from 635 nm to 1.55 μm [12].

In Reference [13], a linear-array spatial light modulator with a linear pixel-array geometry was proposed. The time distribution of femtosecond light pulses was controlled by a computer, and the pixel width was 1000 mm. The existing version handled beams of light with wavelengths ranging from 450 to 2000 nm. Experiments showed that the array had more than 700 deflection-quantization angles and a response time of less than 200 μs [13].

Reference [14] proposed a two-dimensional liquid-crystal spatial light modulator with an effective light-passing aperture of $15.36 \times 8.64\ \text{mm}^2$. It had a unit element size of 8 μm, operating-wavelength range of 420–1700 nm, diffraction efficiency of 93%, and response time of 5.6 ms [14].

In 2022, Meadowlark Corporation of the United States launched a pure-phase liquid-crystal spatial light modulator with a high liquid-crystal response speed (2 kHz at 532 nm), high light-utilization efficiency (98%), high fill factor (97.2%), and high resolution (1024×1024). A silicon-based liquid-crystal spatial light modulator was specially designed for pure-phase applications and combined analog data addressing with a high refresh rate. This combination provided users with the fastest response time and high phase stability.

Domestic Research Progress

In 1994, Huang et al. proposed a new type of liquid-crystal light-valve (LCLV) projection system based on amorphous silicon (a-Si). The system had a short

response time, wide input spectrum, and simple manufacturing process. Experiments showed that the display ratio of the system was 1.5 × 1.2 m^2, the illuminance was 190 lx, the response time was less than 1 ms, and the resolution was above 20 lp/mm [15].

In 1995, Han manufactured an nc-Si–H/a-Si–H liquid-crystal light valve with a heterogeneous structure. The liquid-crystal light valve used multiple absorption layers. The multilayer absorption layer had a high absorption capacity in the visible and near-infrared ranges, ensuring the high sensitivity of the liquid-crystal light valve [16].

In 2002, Zhao et al. modulated the intensity distribution of spatial light beams through a liquid-crystal spatial light modulator, and obtained a light beam with an output near-field modulation of 1.5: 1. Auxiliary features were added to the liquid-crystal spatial light modulator to adjust the optical path and rotate the light spot. A thin-film transistor liquid-crystal light valve was used, which was an electrically addressable transmission-type working mode. After perfecting the system, it was applied to a high-power laser device [17].

In 2005, Ding et al. introduced a new method of light-beam shaping using a liquid-crystal light valve, studied the incident phase-modulation and spatial transmission characteristics of the shaped light beam, and experimentally concluded that the light valve had an excellent shaping effect [19].

In Reference [18], a liquid-crystal phased array with 1920 × 1080 independent phased-array elements was proposed. The phased array had a diffraction efficiency of 60%, 60-Hz frame rate, and 8-μm array unit element size [18].

In 2008, Zhang et al. developed a liquid-crystal spatial light modulator with 2014 independent phase-control units. The array size was 10 μm, and the thickness of the liquid-crystal cell was 5.8 μm. The working voltage was provided to the liquid-crystal cell through a field-programmable gate array, with more than 60 deflection angles. The maximum deflection angle was 2.0014° [19].

In 2012, Huang et al. developed an optically addressable liquid-crystal spatial light modulator. The experiment tested its performance under continuous light with a 1053-nm wavelength, and showed that it had a maximum transmittance of 85% [20]. In 2016, Huang et al. reported the latest progress in the development of liquid-crystal spatial light modulators and conducted preshielding experiments on high-power laser systems [21].

In 2020, Wang et al. developed a liquid-crystal optical phased array with a large aperture. This liquid-crystal phased array achieved a light-beam deflection of 0–3°, with a diffraction efficiency of more than 80%, a pointing error of less than 90 μrad, and a circuit refresh rate of 150 kHz [22] (Fig. 28.1).

In 2019, Li [23] developed a transmissive liquid-crystal spatial light modulator, which had a unit liquid-crystal layer thickness of 1.5 μm, a diffraction efficiency of 35%, and deflected light beams from 0° to 11° with a resolution of 877 lp/mm.

At the end of 2020, Zhongke MSI realized an autonomous core device through the independent design of a liquid-crystal light-valve device. In 2022, Zhongke MSI launched two self-designed spatial light modulators, namely, 2k-color amplitude-type and 2k-color phase-type spatial light modulators. The technical parameters of the products are listed in Table 28.1.

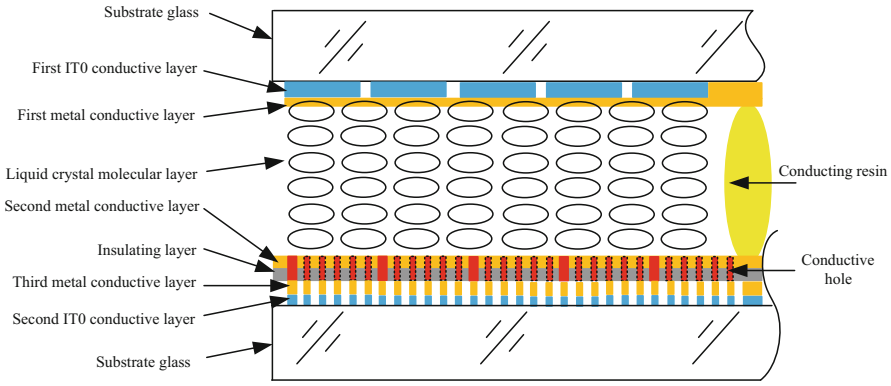


Fig. 28.1 Large-aperture liquid-crystal optical phased-array structure diagram [22]

Table 28.1 Technical parameters of the space light modulator independently designed by Zhongke MSI

Model number	FSLM-2K39-A02	FSLM-2K39-P02
Modulation type	Amplitude type	Phase type
Display type	Monochrome display, synchronous support for color display	
Gray-release grade	256	
Resolution	1920 × 1080	
Pixel size	4.5 μm	
Effective area	8.64 × 4.85 mm^2	
Fill factor	> 91%	
Contrast ratio	≥ 600: 1	–
Phase-modulation range	–	2π (max 3.8π)
Optical utilization	75% ±5%	
Refresh frequency	Monochrome 200 Hz	Color 60 Hz
Spectral range	420–650 nm	
Data interface	Mini DP	
Power supply	DC 5 V, 2 A	

LC-SLM Basic Working Principle

Electronically Controlled Birefringence Properties of Liquid Crystals

In existing research, liquid-crystal spatial light modulators (LC-SLMs) can be divided into two types: twisted nematic spatial light modulators (TN-SLM) and parallel-arrangement nematic spatial light modulators (PA-SLM).

The biggest feature of the TN-SLM is that it has a twisting effect; that is, when a beam of light passes through the liquid-crystal cell, the polarization direction of the light beam is affected by the change in the optical-axis direction of the liquid-crystal

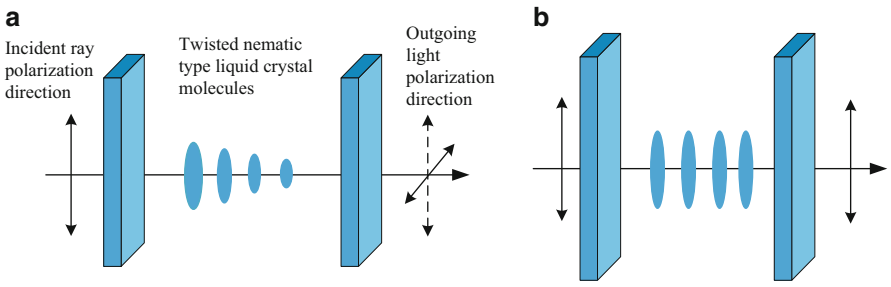


Fig. 28.2 Working principle of a TN-SLM. (**a**) Driving voltage equal to 0; (**b**) driving voltage not equal to 0

cell. As shown in Fig. 28.2, when the left and right surfaces of the liquid-crystal cell are anchored, the liquid-crystal molecules close to them are arranged in a preset direction. The difference between the liquid-crystal molecules on the two sides is $\pi/2$ [24]. The liquid-crystal layer of the twisted nematic spatial light modulator can be regarded as a multilayer stack containing microtwisted uniaxial crystals [25].

Figure 28.2 shows the modulation schematic diagram of a TN-SLM, in which Fig. 28.2a represents the arrangement of liquid-crystal molecules when no external electric field is applied. It appears to be helical at this time. In the absence of an applied voltage, the liquid-crystal layer appears to be weakly twisted. If the optical-axis direction and polarization direction of the incident ray are the same, a change in the former will lead to a change in the latter. The two are closely correlated, and the polarization direction of the outgoing light remains the same as that of the optical axis.

At this time, it can be considered that the phase difference between the o-light and e-light is fixed, and its expression is as follows:

$$\phi = \frac{\pi d(n_e - n_o)}{\lambda}, \tag{28.1}$$

where d represents the thickness of the liquid-crystal layer and λ is the wavelength of the modulated light wave.

As shown in Fig. 28.2b, the presence of an electric field has a significant effect on the liquid-crystal molecules, resulting in changes in the birefringence of the liquid crystal. This phenomenon occurs because the degree of electric-field influence on each position of the liquid-crystal cell is different. The surface is basically unaffected, whereas the interior, especially the middle part, is more susceptible. Therefore, when the voltage is higher, the number of molecules arranged in the direction of the electric field increases significantly, reducing the distortion effect to a certain extent.

In addition, when the rotation angle q changes, n_e will change accordingly; however, it does not affect the size of n_o. The phase difference ϕ between the ordinary and extraordinary light will also change according to the change in the rotation angle. At this point, the equivalent refractive index of extraordinary light $n_e(\theta)$ can be expressed as follows:

$$\frac{1}{{n_e}^2(\theta)} = \frac{\cos^2\theta}{{n_e}^2} + \frac{\sin^2\theta}{{n_o}^2}. \tag{28.2}$$

In the formula, the size of $n_e(\theta)$ is related to the characteristics of the liquid crystal. When the TN-SLM modulates the phase, the amount of modulation is directly related to the applied voltage. In the phase-modulation process, the amplitude is modulated to a certain extent; therefore, to modulate the amplitude, a polarizer should be added to the system.

Figure 28.3 shows the nematic liquid-crystal spatial light modulator arranged in parallel, and the liquid-crystal molecules arranged in parallel. The liquid-crystal molecules at the junction of the electrode surface are always fixed and maintain a vertical relationship with the electrode surface, which does not twist with the change in the applied electric field [26].

Figure 28.3a shows that when there is no electric-field force, the liquid-crystal molecules do not have a twisting effect, and the molecules are parallel to each other and to the liquid-crystal cell. Figure 28.3b shows that when an electric field is applied, the liquid-crystal molecules rotate such that the optical-axis direction of the molecules is perpendicular to the direction of the applied electric field. This is because the liquid-crystal molecules have an elastic force, and under the action of the electric-field force and molecular-elastic force, the liquid-crystal molecules rotate [27, 28].

When a voltage V_r is applied to the PA-SLM, the rotation angle of the liquid-crystal molecules θ can be described by the magnitude of the voltage as [24]

$$\theta = \begin{cases} 0 & , V_{rms} < V_c \\ \dfrac{\pi}{2} - 2\tan^{-1}\left\{\exp\left[-\left(\dfrac{V_r - V_c}{V_o}\right)\right]\right\}, & V_{rms} \geq V_c \end{cases}, \tag{28.3}$$

where V_c and V_o represent the threshold voltage and saturation voltage, respectively. It can be seen from Eq. 28.3 that the rotation conditions of the liquid-crystal molecules are as follows: When $V_r < V_c$, the liquid-crystal molecules deflect, and when $V_r > V_c$, the liquid-crystal molecules do not rotate.

When liquid-crystal molecules are subjected to an electric-field force, the changes in n_e and n_o can be obtained by analyzing the refractive-index ellipsoid, from which

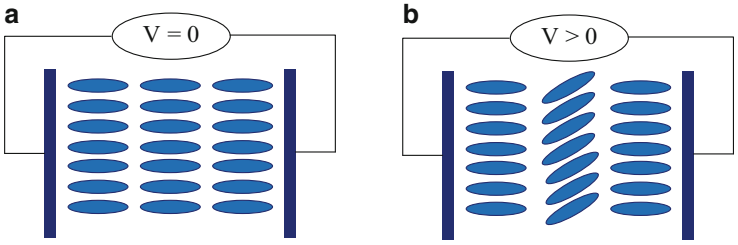


Fig. 28.3 Working principle of a PA-SLM [29]. (**a**) Driving voltage equal to 0; (**b**) driving voltage not equal to 0

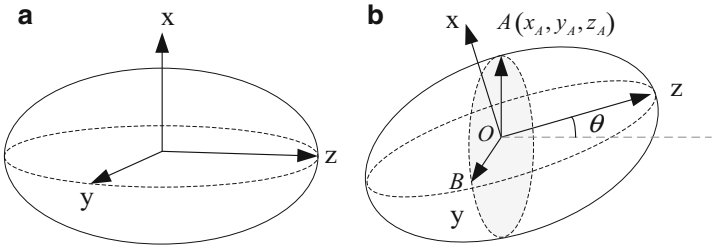


Fig. 28.4 Refractive-index ellipsoid. (**a**) Applied voltage is equal to 0; (**b**) applied voltage is not equal to 0

the liquid-crystal phase-transition principle can be deduced [29]. When there is no external electric-field force, the direction of the long axis changes and the direction and tendency of the electric field maintain a vertical relationship. Influenced by the electric-field force and elastic force of the liquid-crystal molecules, the liquid-crystal molecules begin to rotate. Its angle is expressed as θ, and the result is shown in Fig. 28.4b [23].

In Fig. 28.4, the long axis of the liquid-crystal molecule, O_z, selects a point on the ellipsoid, A (coordinates (x_A, y_A, z_A)); the origin of the ellipse is O; then, the expression is obtained as

$$\frac{x_A^2}{n_1^2} + \frac{z_A^2}{n_2^2} = 1. \tag{28.4}$$

In the formula, $x_A = OA \cos\theta$, $z_A = OA \sin\theta$, $OA = n_e$, the intersection of the cross section of points O and A and the y-axis is represented by B, and the value of OB is n_o. The short and long semi-axes of the ellipse are n_1 and n_2. Therefore, if the liquid-crystal molecules rotate, the relationship between n_e and n_o is as follows:

$$\begin{cases} \dfrac{1}{n_e^2(\theta)} = \dfrac{\cos^2\theta}{n_1^2} + \dfrac{\sin^2\theta}{n_2^2} \\ n_0 = n_1 \end{cases}. \tag{28.5}$$

Combined with the information in Fig. 28.4, it can be seen that the rotation process of the liquid-crystal molecules has certain particularities. When the incident ray is parallel to the surface, the phase-modulation process can be completed, owing to the change in the OA length. For incident rays that remain perpendicular to the surface, the OB length is fixed; therefore, the relationship between the polarization direction and the surface is also fixed and will not change. Therefore, it can be concluded that the direction of the polarization-state component affects the modulation effect.

If the incident ray is perpendicular to the surface, there is no modulation effect; however, when it is parallel, there is a modulation effect. During the phase-modulation experiment, it was necessary to ensure that there were no polarization

components in other directions, so that the liquid-crystal spatial light modulator could work in a pure phase state without affecting the intensity. A polarizer or half-wave plate was added to the experimental optical path.

It can be seen from the above analysis that when the polarization direction is parallel to the optical-axis direction, the polarization direction is fixed. At this time, only the phase of the outgoing light changes, and pure phase modulation can be performed. When the molecular optical axis is not parallel to the polarization direction, the polarization direction and the phase of the outgoing light change.

Directional Distribution of Liquid Crystals

The pointing vector of liquid-crystal molecules is used to characterize the direction vector of the overall liquid crystal. The pointing average of the long sides of the molecules is a statistical average of the local part. The pointing vector of liquid-crystal molecules can be analyzed using the Cartesian coordinate system, as shown in the figure, where φ is the azimuth angle and $\theta = \theta(z)$ is the inclination angle [25].

As shown in Fig. 28.5, the unit vector can be expressed as

$$\vec{n} = (\cos\theta\cos\varphi,\ \cos\theta\sin\varphi,\ \sin\theta). \tag{28.6}$$

According to the elastic-deformation theory of the liquid-crystal continuum, the Gibbs free energy in a liquid-crystal cell is composed of splay, twist, and bend deformations [30]:

$$W_g = \int_0^d f_g dz = \frac{1}{2}\int_0^d \left\{ K_S\left(\nabla\cdot\vec{n}\right)^2 + K_T\left(\vec{n}\cdot\nabla\vec{n}\right)^2 + K_B\left(\vec{n}\times\nabla\times\vec{n}\right) - \vec{D}\cdot\vec{E}\right\} dz, \tag{28.7}$$

where K_S is the S coefficient, K_T is the T coefficient, K_B is the B coefficient, and ε is the dielectric tensor that constitutes the liquid-crystal cell material. The electric-field strength is expressed in $\vec{E}$ and the electrical displacement $\vec{D} = \varepsilon\vec{E}$. The free-energy density can be expressed as follows:

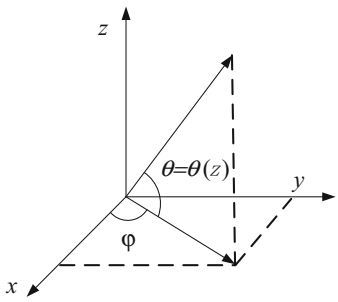


Fig. 28.5 Schematic diagram of the vector distribution of liquid-crystal molecules

$$f_g = \frac{1}{2}\left[f(\theta)\left(\frac{d\theta}{dz}\right)^2\right] - \frac{1}{2}\left(\frac{dU}{dz}\right)^2\left(\varepsilon_{\parallel}\cos^2\theta + \varepsilon_{\perp}\sin^2\theta\right), \tag{28.8}$$

where $\varepsilon_{\parallel}$ and $\varepsilon_{\perp}$ respectively denote parallel and perpendicular to the direction vector:

$$f(\theta) = K_S\cos^2\theta + K_B\sin^2\theta$$

$$g(\theta) = \left(K_T\cos^2\theta + K_B\sin^2\theta\right)\cos^2\theta.$$

By calculating the partial differential of $f(g)$ with respect to the inclination angle θ, azimuth angle φ, and potential U, the following equations can be deduced:

$$\begin{aligned}
&\frac{\partial f_g}{\partial\theta} - \frac{d}{dz}\left[\frac{\partial f_g}{\partial\left(\frac{d\theta}{dz}\right)}\right] = 0\\
&\frac{\partial f_g}{\partial\varphi} - \frac{d}{dz}\left[\frac{\partial f_g}{\partial\left(\frac{d\varphi}{dz}\right)}\right] = 0\\
&\frac{\partial f_g}{\partial U} - \frac{d}{dz}\left[\frac{\partial f_g}{\partial\left(\frac{dU}{dz}\right)}\right] = 0.
\end{aligned} \tag{28.9}$$

By solving the differential equation of Eq. 28.9 and substituting the boundary conditions, the vector distribution can be obtained.

Reasons for and Experiments with the Phase Calibration of Liquid-Crystal Spatial Light Modulators

Liquid-crystal material is used in the phase modulation of spatial light modulators; however, it is limited by the liquid-crystal material, resulting in several problems in the spatial light modulator that cannot be ignored.

1. Owing to the different photoelectric characteristics of the individual liquid-crystal pixels in the liquid-crystal spatial light modulator, there is a small deviation in the light-modulation process, resulting in an uneven phase modulation in the entire space. This inhomogeneity is related to the surface-protection glass of the spatial light modulator and the curvature of the baseplate of the address-selection circuit. The curvature of the backplate of the address-selection circuit is the result of the different polishing rates of the nonmetallic and metal regions of the backplate

during the polishing process, resulting in an initial phase difference of the liquid-crystal spatial light modulator [31, 32].
2. In the control software of the liquid-crystal spatial light modulator, the initial optimal phase is affected by external factors, which leads to the need to compensate for the initial deviation, thereby increasing the error [33, 34].
3. The dispersion effect of liquid-crystal molecules [35] causes the device to have a 2π phase modulation for a single set wavelength [36].

Therefore, before actual use, it is necessary to calibrate the liquid-crystal spatial light modulator to achieve a better correction effect [37]. At present, under different-wavelength light beams, the phase-calibration methods of liquid-crystal spatial light modulators include interferometry and diffraction methods [38, 39]. Interferometry mainly includes double-slit interference, Mach–Zehnder interference [40, 41], and Twyman–Green interference [42, 43]. Interferometry primarily uses the relative displacement of the fringes for phase calibration. The two beams of light pass through a long path before interference. Under the interference of external factors, the optical-path difference is affected, and unstable fringes are collected [44]. The following two interferometric methods are introduced.

Double-Slit Interference

Double-slit interferometry is shown in Fig. 28.6. The specific implementation process is as follows: When the light beam is punched into two sets of double slits, it generates two sets of interference fringes, and the distance between these is reasonably fixed. Relative movement will occur and, using this relative-movement amount, the relationship between the phase-modulation amount of the liquid-crystal spatial light modulator and the added gray release can be calculated.

This method requires a reference group, and the only phase modulation obtained by this method is the phase information on the phase seam; thus, the phase-modulation characteristics of the liquid-crystal spatial light regulator cannot be derived [45].

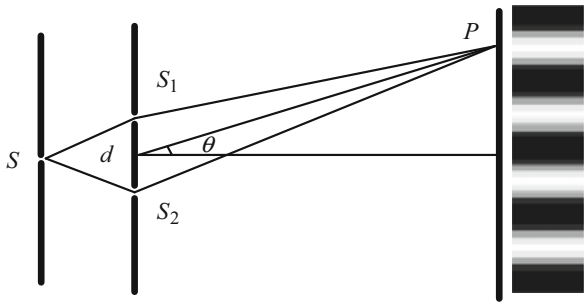


Fig. 28.6 Schematic diagram of double-slit interference

Twyman–Green Interference (Fig. 28.7)

The specific implementation process of Twyman–Green interferometry is as follows: A beam-splitting prism is used to split the incident ray. After the beam is split, one of the light beams is reflected by a mirror, and the other beam is reflected by a liquid-crystal spatial light modulator. When the two beams are reflected back, a splitting prism is used to interfere with them, and a charge-coupled device (CCD) is used to record the interference fringes.

During the experiment, it was necessary to load the gray-release images of each order on the LC-SLM so that the interference fringes obtained by the CCD camera could be relatively displaced. The relationship between the gray-release value and phase delay can be determined according to the size of the Orvin shift [46, 47]. Through numerical calculations and experimental calibration, it was concluded that this interferometry was a good calibration of the phase of the reflective liquid-crystal spatial light regulator. Mach–Zehnder and shear interferometry were used for the phase measurement and calibration of the transmissive liquid-crystal spatial light regulator [48–51].

Turbulence-Simulation Principle of a Liquid-Crystal Spatial Light Modulator

The atmospheric refractive index is affected by atmospheric turbulence [52]; therefore, when a beam of light travels in the atmosphere, it undergoes phase distortion, light-intensity flickering, light-spot drift, and other effects. The commonly used methods for simulating atmospheric turbulence cannot accurately simulate the phase, spatial, or temporal characteristics of atmospheric turbulence. Therefore, it

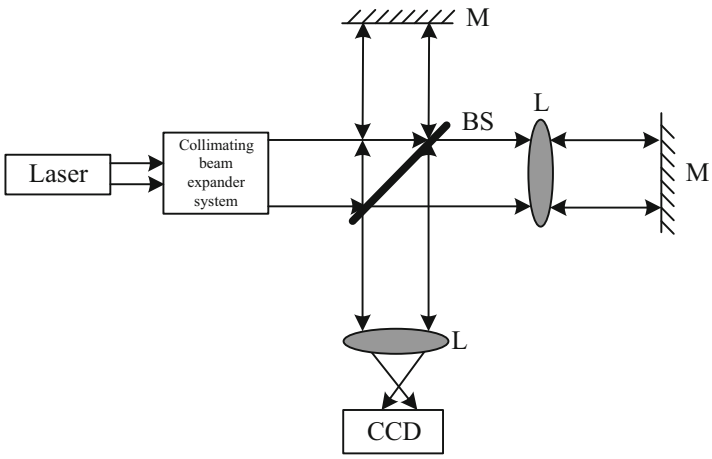


Fig. 28.7 Schematic diagram of Twyman–Green interference

is necessary to use suitable devices to simulate atmospheric turbulence in a laboratory under natural conditions. Thus, not only can the influence of atmospheric turbulence on light waves be studied, but the stability and reliability of the adaptive-optical system can be verified, as well as studying the physical properties of atmospheric turbulence [53].

According to the frozen-turbulence hypothesis [54], it can be concluded that as a beam propagates in atmospheric turbulence, random phase screens can be used to simulate the distortion effect. Therefore, related studies can be conducted using numerical simulations. Commonly used methods can be divided into two categories: indirect simulations in the frequency domain and direct simulations in the spatial domain. Their specific methods differ.

For the former, the power-spectrum inversion method (also known as the Fourier-transform method) is mainly used, and for the latter, the Zernike polynomial method is used. In addition, there are fractal methods [55] and covariance-function methods [56]. The simulation of the atmospheric-transport turbulence effect is mainly divided into two stages: a numerical simulation of the turbulence description and a physical representation of the hardware-in-the-loop simulation.

The specific process of the hardware-in-the-loop simulation is as follows: Through the liquid-crystal spatial light modulator, the phase of the laser wavefront is modulated, and the random energy distribution of the laser in time and space is simulated. Subsequently, the randomly generated phase data sequence is processed and injected into the liquid-crystal spatial light modulator, which modulates the laser. The random phase and far-field distribution of the intensity fluctuation of the laser light spot can be obtained.

An electronically addressable liquid-crystal spatial light modulator was used to complete the hardware-in-the-loop simulation of the turbulence. The main idea is to modulate the phase of the outgoing light using a spatial light modulator [57]. The wavefront can be modulated by changing the voltage alignment of the liquid-crystal molecules, when they form a specific equivalent-phase plane inside the liquid-crystal cell [58].

When there is an electric field in the optical-axis direction, the liquid-crystal molecules in the middle of the substrate deflect. At this time, the equivalent refractive index n_{eff} of the liquid-crystal molecules can be obtained. It is related to the angle θ between the long axis and the x-axis of the liquid-crystal molecules, that is, [59]

$$n_{eff} = \frac{n_o n_e}{\sqrt{n_o^2 \cos^2\theta + n_e^2 \sin^2\theta}}. \tag{28.10}$$

After linearly polarized light with wavelength λ passes through the liquid-crystal layer with thickness d, its phase difference $\Delta\varphi$ can be expressed as

$$\Delta\varphi = \frac{2\pi}{\lambda} \int_{-d/2}^{d/2} \left[n_{eff}(\theta) - n_o \right] dz. \tag{28.11}$$

After simulating the turbulence phase change, the light intensity obtained on the detector surface is

$$I_d = I_0 + I_1 \sin \varphi_e, \tag{28.12}$$

where I_0 is the background light, $I_1 = \mathrm{C}\,|E_m{}^2|$ is the modulated light intensity, $\mathrm{C_o}$ is the amplification coefficient, and $|E_m{}^2|$ is the amplitude of the light emitted by the spatial light modulator.

$$E_m = |E_i(x, y, t)| \times \exp j[\phi_i(x, y, t) - \phi_m(x, y, t) - \phi_o(x, y)], \tag{28.13}$$

where φ_i and φ_m are the phases of the input spatial light modulator and the output signal of the spatial light modulator, respectively.

(x, y, t) is used to describe the time-related position parameters on the light-modulation surface of the liquid-crystal spatial light modulator, and φ_o is used to represent the static phase error caused by the system and the external environment of the liquid-crystal spatial light modulator. In the case of an actual turbulence simulation, a voltage phase lookup table of the liquid-crystal spatial light modulator is mainly used, and the electrode voltages of different pixels are controlled by programming to modulate the phase of the wavefront and simulate the turbulence [60].

Wavefront-Distortion Correction Principle of a Liquid-Crystal Spatial Light Modulator

The static wavefront-correction principle of a liquid-crystal spatial light modulator is based on the phase-conjugation principle. The phase-conjugation principle is equivalent to that of a photoelectric-field operator. The characteristic of a phase-conjugation light is that the beam of light is disturbed by uneven media, and the interference when the phase-conjugate light passes through the medium is cancelled. This is because the phase-conjugation principle is widely used to correct the phase distortion of light beams [61].

According to the wavefront-conjugation principle, the complex amplitude of the received wave can be conjugated to compensate for the wavefront distortion experienced by the light wave during transmission. Suppose that the complex-amplitude expression of the received wave is

$$\psi(r, t) = |\psi(r, t)| \exp[i\phi(r, t)], \tag{28.14}$$

where ψ (r, t) and φ (r, t) are used to express the amplitude and phase of the received wave, respectively, $r = \{x, y\}$ is the position vector perpendicular to the direction of light propagation, r is the magnitude, and t is the time. Through the principle of wavefront-conjugate correction, the complex amplitude of the emitted wave can be expressed as

$$A(r, t) = \alpha\psi^*(r, t) = \alpha|\psi(r, t)| \exp[-i\phi(r, t)] = |A(r, t)| \exp[iu(r, t)], \tag{28.15}$$

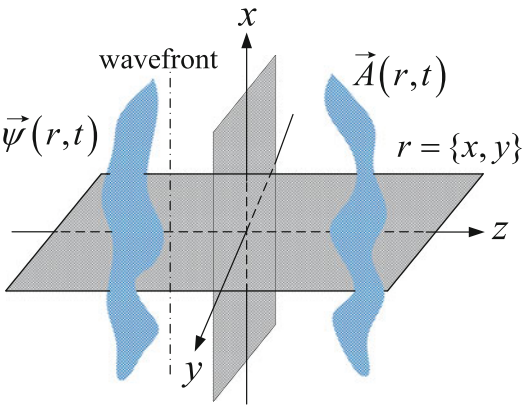


Fig. 28.8 Schematic diagram of wavefront conjugation

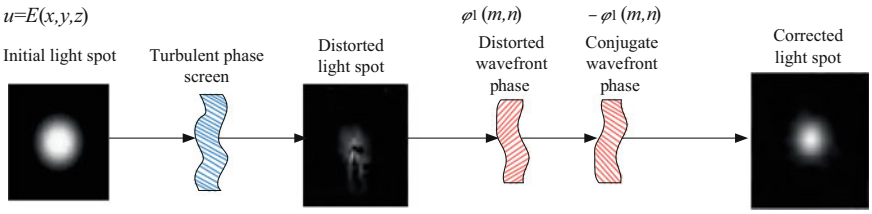


Fig. 28.9 Schematic diagram of static wavefront-conjugate correction

where $|A(r, t)|$ and $u(r, t)$ represent the amplitude and phase after light-wave correction, respectively, and α is the real coefficient. Therefore, the phase conjugate $u(r, t) = -\phi(r, t)$ and amplitude modulation $|A(r, t)| = \alpha|\psi(r, t)|$ constitute the wavefront-conjugate conditions [62]. If the light wave travels along the z-axis, a wavefront-conjugate schematic can be obtained, as shown in Fig. 28.8. A schematic of the experimental setup is shown in Fig. 28.9 [63].

Classification of Liquid-Crystal Spatial Light Modulators

In Fig. 28.10, transmissive-type spatial light modulators are shown in Fig. 28.10a, c, and reflective-type spatial light modulators are shown in Fig. 28.10b, d. Moreover, electrically addressable-type spatial light modulators are shown in Fig. 28.10a, b, and light-addressable-type spatial light modulators are shown in Fig. 28.10c, d.

The photosensitive surface of the spatial light modulator is an important structure in the optical-addressing implementation process. When the two-dimensional light-intensity distribution is projected onto it, each individual pixel of the spatial light modulator will obtain a corresponding optical signal, and through these optical signals, find the address. Similar to optical addressing, electrical-addressing spatial

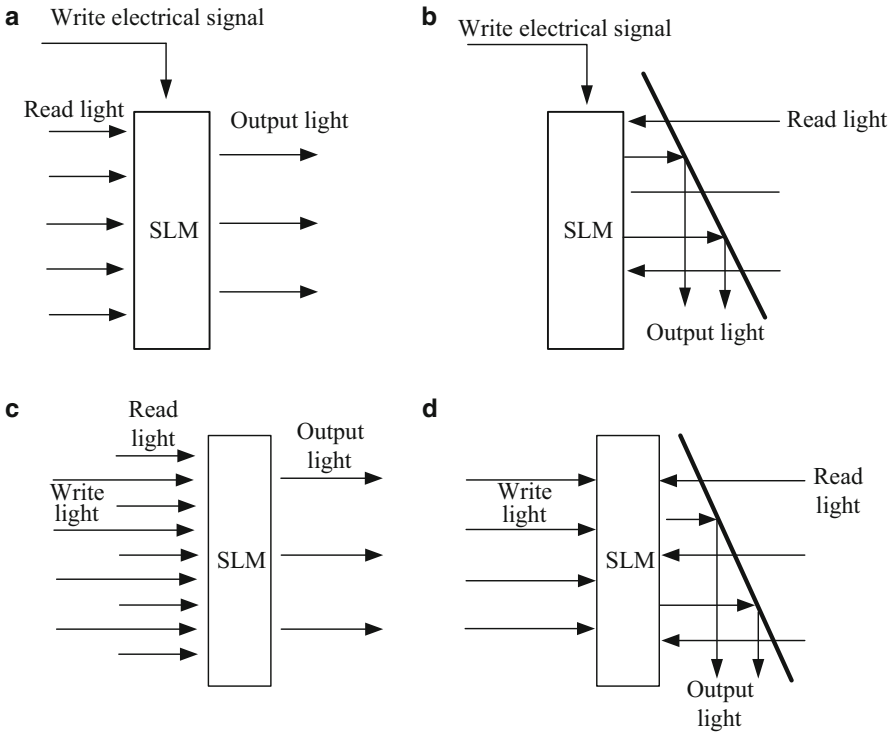


Fig. 28.10 Schematic diagram of four spatial light modulators [62]. (**a**) Electrically addressed transmissive; (**b**) electrically addressed reflective; (**c**) optically addressed transmissive; (**d**) optically addressed reflective

light modulators implement addressing by applying a voltage to each pixel through a circuit [64].

Liquid-crystal spatial light modulators are classified according to the modulation parameters of the light wave. If the amplitude of the light wave is modulated, it is called an amplitude-type spatial light modulator, which primarily modulates the light intensity of incident rays. A phase-type spatial light modulator primarily modulates the wavefront phase of the light wave and is often used to compensate and correct the front of the light wave [65].

Research Status of Optical-Modulation Technology

Foreign Research Progress

In the late twentieth century, McManamon et al. investigated the optical phase-modulation process and concluded that optical-phased arrays had the potential to overcome the limitations of mechanical light-control beams. Lindquist et al. reported

that liquid-crystal phase gratings had a high diffraction efficiency and strong incident-angle dependence. The high anisotropy of the grating demonstrated the properties of blazed gratings, which made the gratings useful for light-beam control and diffraction applications [66, 67].

In 1997, Titus et al. studied the mathematical model of a liquid-crystal cell and used it to construct a diffraction grating. They designed a pure phase grating whose phase distribution was easy to control, and analyzed the change in pointing accuracy [68].

In 2001, Hakamura et al. studied the effect of the overdrive addressing scheme on a liquid-crystal display by applying an initial pulse slightly higher than the rated voltage to the target; it caused the fast director to redirect, depending on the dynamics. The results showed that the overdrive method greatly reduced the liquid-crystal response time and improved the blurred tailing of the image [69].

In 2003, Harris et al. described a simple experimental setup that could be used to measure the steering efficiency of liquid-crystal optical phased arrays. Several optimization methods were discussed to determine the voltage per electrode to optimize the steering efficiency. The experimental results showed that this method could quickly determine the optimal voltage for an ideal far-field diffraction pattern. Compared with the calibration of the open-loop device, the steering efficiency was improved by more than 100% [70].

In 2004, McManamon studied liquid-crystal optical phased arrays through numerical simulations, discussed the effects of the edge electric field, liquid-crystal material properties, and voltage-optimization process on the performance of liquid-crystal optical phased arrays, and discussed the structure of liquid-crystal optical phased arrays. The diffraction angle and diffraction efficiency are related design issues. They proposed a wide-angle design of liquid-crystal optical phased arrays for efficient laser-beam control [71].

In 2004, Barrick et al. proposed an electronically controlled birefringence device that controlled a beam of light using components approximately 2–20 times the wavelength of light inside a liquid crystal. These components were narrow transparent electrodes that formed a sagittal profile along one dimension of the liquid-crystal plate, thereby affecting the phase of the coherent light passing through the device. When dual-frequency liquid-crystal materials were exposed to applied voltages of different frequencies, they exhibited dielectric anisotropy of opposite signs. By using high frequencies to drive liquid-crystal molecules to a state close to their zero-voltage equilibrium conformation, the material improved the switching speed of the device. In the experimental verification, the model for 1550-nm lasers was used; when the beam of light was deflected by 0.86–8.86°, the optimized beam-diffraction efficiency was significantly improved [72].

In 2006, Tholl et al. discovered that a beam of light could be steered by applying a linear phase delay to the aperture of the laser beam. The corresponding wavefront slope determined the steering angle. The combination of a large wavefront slope and large aperture required a large optical-path difference across the aperture, which needed to be achieved by the beam of a light-control device. Based on this, a new pointing-accuracy optimization method was developed [73].

In 2006, Robert et al. developed a programmable circuit for fast-switching liquid-crystal spatial light modulators. It was programmed to improve the speed of nematic liquid-crystal spatial light modulators with a very slow response speed, reducing the modulation phase and the time required. The response time required for the phase to rise from 0 to 2π was shortened from 320 ms to 67 ms, and the response time required for the phase to rise from 2π to 0 was shortened from 420 ms to 8 ms [74].

In 2008, Engstrom et al. found that when using a one-dimensional phase spatial light modulator to control a beam of light, the limited pixel and its quantized phase-modulation value limited the modulation accuracy. Therefore, a method of modulating the phase surface was proposed to calibrate the front of the light wavefront, and the pointing accuracy was optimized by determining the average slope of the wavefront after calibration [75].

In 2011, Kim et al. described and demonstrated a nonmechanical light-beam control device based on passive gratings and liquid-crystal polymers, which achieved a large-angle deflection and had high efficiency. With a nonmechanical light-beam steering range of 80°, for a light beam with a 1064-nm wavelength, there were 32 steering angles, and the angular resolution was 2.6°. The system was efficient at 66–70% throughput, but was limited by reflection and absorption losses caused by the electrode materials [76].

In 2013, Thalhammer et al. proposed two mutually compatible methods for accelerating the dynamic response speed without a loss of diffraction efficiency. One technique involved an overspeed drive, which calculated intermediate modes to expedite the transition to target phase modes. The second method optimized the target pattern to reduce the desired phase change applied to each pixel, which resulted in a dramatic reduction in the change in diffracted-light intensity during the transition. When these methods were applied together, the transition time of the diffracted light field was approximately 1 ms, which was a tenfold improvement over the current methods. Experiments showed that the method was improved in holographic-image projection, light-beam control and switching, and real-time control loops [77].

In 2017, Steven et al. developed a 1550-nm light-beam deflection system with a net aperture of 10 cm that could be steered at a 1° resolution at a 64° × 64° scanning field of view. The light-beam-diffraction efficiency reached more than 80% and maintained the wavefront distortion at $\lambda/10$ root-mean-squared (RMS) [78].

In 2020, Xiao et al. developed a diffraction grating with a large scanning field of view and high diffraction efficiency with a 532-nm wavelength and 400-nm period transmission mode. The angular bandwidth was significantly increased by arranging two inclined grating layers in the same monolithic film. By simulating the optical properties, a diffraction structure with a 48° scanning field of view and diffraction efficiency of 70% was determined [79].

In 2019, He et al. proposed the use of several phase-recovery techniques to directly measure the corresponding wavefront and used the obtained wavefront to correct aberrations caused by spatial inhomogeneity. Using Twyman–Green interferometry, an interferogram-acquisition technique based on SLM wavefront modulation was proposed. In addition, a method for determining the corresponding active

area of the liquid-crystal backplane and the parameters of the two-dimensional projection transformation was proposed to convert the active area into the fitting resolution of the spatial light modulator [80] (Table 28.2).

Table 28.2 Research status of foreign key optical-modulation technologies

Year	Researcher	Key technology	Research results
End of twentieth century	McManamon	Optical phase-modulation process	Nonmechanical light-beam control using an optical phased array
End of twentieth century	Lindquist	Characteristics of liquid-crystal phase grating	Liquid-crystal grating for light-beam control and diffraction applications
1997	Titus	Mathematical model of a liquid-crystal cell	A pure phase grating with easy phase-distribution control was designed
2001	Hakamura	Overdrive addressing scheme	Overdrive method reduced the liquid-crystal response time
2003	Harris	Measured the steering efficiency of a liquid-crystal optical phased array	Proposed a method to optimize steering efficiency
2004	McManamon	Numerical simulation of a liquid-crystal optical phased array	Proposed a wide-angle liquid-crystal optical phased array with efficient laser-beam control
2004	Barrick	Electronically controlled birefringent liquid-crystal phased array	Optimized light-beam-diffraction efficiency
2006	Tholl	Relationship between light-beam steering and liquid-crystal aperture	Researched a new pointing-accuracy optimization method
2008	Engstrom	Influence of number of pixels and phase quantization on the modulation accuracy	Proposed a method of modulating the phase surface to calibrate the light wavefront
2011	Kim	Light-beam control using a passive grating and liquid-crystal polymer	Improved the deflection efficiency and detailed quantification of the deflection angle
2013	Thalhammer	Method to speed up the dynamic response of liquid crystals	Holographic image projection, light-beam control, and switching and real-time control loops
2017	Steven	1550-nm light-beam deflection system	Improved the diffraction efficiency of a light beam and stabilized the wavefront distortion
2020	Xiang	New diffraction grating	Improved the scanning field of view and diffraction efficiency
2019	He	Phase-recovery technology to directly measure the corresponding wavefront	Interferogram-acquisition technique based on SLM wavefront modulation

Domestic Research Status

In 2009, Kong et al. proposed a method for quantitatively calculating the scanning error and found that it was on the order of 10^{-1}, 10^{-2}, and 10^{-3}. They investigated the influence of the manufacturing process of a lidar liquid-crystal phased array element on the scanning accuracy [81, 82].

In 2011, Zou cascaded a multiplex Bragg grating with a liquid-crystal spatial light modulator. The system controlled the light beam to achieve quasi-continuous deflection at 0–11.82°. The cascaded system had a maximum deflection efficiency of approximately 2% for the beam of light [83].

In 2011, Yin built a closed-loop optimization experimental platform and proposed a diffraction-efficiency optimization method based on energy constraints that did not require a probe wavefront phase. The phase-recovery method was used to correct the phase difference between array elements caused by the inhomogeneity of periodic-array elements, which effectively improved the diffraction efficiency of the light beams [84].

In 2012, Zhou et al. developed an optimization algorithm that used the slope-correction accuracy of a single electrode. This method adjusted the gap between the approximate slope and the ideal slope by narrowing the voltage step height and optimized the pointing accuracy [85].

In 2014, Qiao established a lidar system using one-dimensional liquid-crystal optical phased-array scanning, which laid the foundation for an analysis of lidar accuracy. The voltage-quantization noise caused by the side lobes and extension of the laser pulses were analyzed. A scanning-imaging lidar-system platform based on a liquid-crystal optical phased array was established, and steering and ranging accuracy experiments were performed. An iterative phase-compensation method was used to improve the steering accuracy, and the side-lobe suppression method was used to improve the ranging accuracy. The simulation results demonstrated that the proposed method was feasible. When the pointing accuracy $n = 16$, the steering accuracy improved from 0.13 mrad to 0.04 mrad, and the ranging accuracy improved from 0.040 m to 0.021 m [86].

In 2017, Yang et al. built a test system with a silicon-based liquid-crystal spatial light modulator with 1920 × 1080 pixels to realize the intelligent nonmechanical control of light beams. The system divided the incident ray beam into two, three, four, and five beams to separately control each beam [87].

In 2017, Wang et al. proposed a nonmechanical laser-beam control scheme to achieve high-efficiency continuous two-dimensional scanning. The core method was to orthogonally cascade two liquid-crystal optical phased arrays and use them for azimuth and elevation. The steering characteristics of the two-dimensional light beam were verified experimentally. Based on the experimental device, linearly polarized beams with different polarization angles were controlled with high precision. The measured angular deviation was less than 5 μrad, and the beam-deflection efficiency was 84.7% [63] (Table 28.3).

In 2020, Xiao derived the transformation formula of a Gaussian light beam passing through a telescope system to realize the wide-angle deflection of a liquid-

Table 28.3 Research status of key technologies in domestic light modulation

Year	Researcher	Key technology	Research results
2009	Kong Ling Jiang	Error-scanning method	Proposed the influence of a manufacturing process on precision scanning
2011	Zou Yongchao	Cascaded Bragg grating and liquid-crystal spatial light modulator	Controlled beams of light to achieve quasi-continuous deflection
2011	Yin Xiaohu	Diffraction-efficiency optimization method based on energy constraint	Improved the light-beam diffraction efficiency
2012	Zhou Yan	Optimization algorithm for single-electrode slope-correction accuracy	Optimized the pointing accuracy
2014	Qiao Tianyuan	Liquid-crystal optical phased array LiDAR system	Laid the foundation for a lidar-accuracy analysis.
2017	Yang He	Nonmechanical intelligent control of light beams	Separately controlled multiple single light beams and multiple beams of light
2017	Wang Xiangru	Quadrature cascade of two liquid-crystal optical phased arrays	High-precision control of linearly polarized beams
2020	Xiao Feng	Angle magnification	Deduced the transformation formula of a Gaussian light beam passing through a telescope system
2019	Li Songzhen	Combination of four liquid-crystal polarization gratings	High-precision control of light-beam deflection
2019	Ke Xizheng	Conjugate-wavefront method	Realized static wavefront correction and improved the Strehl ratio
2020	Niu Qifeng	Derived a calculation model of the actual inclination angle and driving voltage	Precisely controlled the deflection of beams of light
2022	Chen Lujian	Microfluidic technology	Real-time continuous modulation of operating wavelengths

crystal optical phased array [88]. The optical transfer matrix of the telescope was studied and the angular-amplification effect was derived.

In 2019, Li combined a liquid-crystal polarization grating with a liquid-crystal phased array to realize a high-angle, high-precision laser scanning system. Four liquid-crystal polarization gratings were combined to achieve a one-dimensional light-beam deflection with a $\pm 20^\circ$ angle range; the grating vector of the same group of liquid-crystal polarization gratings was rotated 90° to realize a two-dimensional large-angle light-beam deflection. The polarization grating was connected to the liquid-crystal phased array by designing the optical path, and the scanning step size of the grating was filled with the liquid-crystal phased array to achieve a high-precision light-beam deflection [89].

In 2019, Ke et al. used a liquid-crystal spatial light modulator to mount an adaptive-optics system using the conjugate-wavefront method to compensate for

the distortion before the wave and realize static wavefront correction. The Strehl ratio was 0.3795 before correction and 0.8268 after correction [62].

In 2020, based on the establishment of an electrical model of the multilayer physical structure of a liquid-crystal spatial light modulator, Niu et al. proposed a quantitative-calculation model for the mapping relationship between the actual inclination angle and the driving voltage, which accurately controlled the light-beam deflection [90].

In 2022, Chen et al. achieved the real-time continuous modulation of operating wavelengths in anisotropic polymerized cholesteric liquid-crystal (PCLC) films with the help of a microfluidic-based "dip-refill" strategy, with a dynamic range of up to 210 nm [91].

Performance Evaluation of Liquid-Crystal Spatial Light Modulators

Amplitude-Modulation Characteristics

With the twisted nematic liquid-crystal characteristics, a liquid-crystal spatial light modulator can change the polarization state of an incident ray and modulate the light amplitude by selecting a suitable polarizer [92]. Figure 28.11 shows the principle of modulating the amplitude of the twisted liquid crystal. The long-axis distortion angle of the liquid-crystal molecules on the positive and negative surfaces of the liquid-crystal cell is 90°, and the polarizers with the same long-axis direction and polarization direction on the front and rear surfaces of the liquid-crystal cell are placed in the front and rear of the liquid-crystal cell, respectively.

Amplitude modulation controls the polarization direction of the outgoing light by applying different external-field voltages to both ends of the liquid-crystal cell, thereby controlling the light intensity [93].

Phase-Modulation Characteristics

A liquid-crystal spatial light modulator can be regarded as a polarizing device because of its polarization-dependent characteristics. Through a Jones matrix, the light-intensity transmittance of the light beam passing through the polarizer, liquid-crystal spatial light modulator, and polarizer can be measured [94, 95].

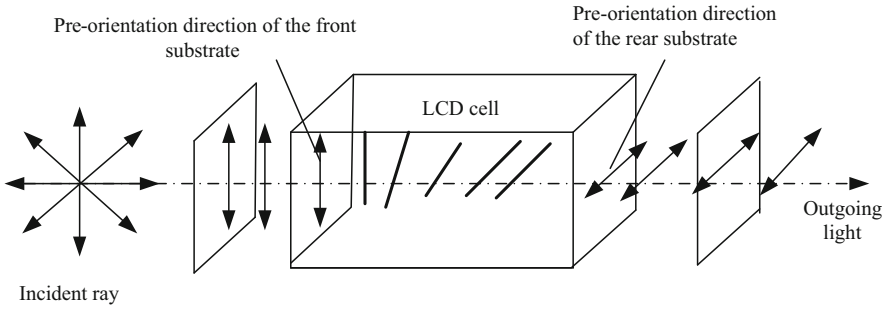


Fig. 28.11 Schematic diagram of liquid-crystal amplitude modulation [93]

Assuming that the incident surface of the liquid-crystal layer is parallel to the x-y plane, the optical axis of the liquid crystal is arranged along the x-axis, and the light propagates along the z-axis. The polarization direction of the polarizer and the x-axis form the angle φ_1, and the polarization direction of the polarizer and the x-axis form angle φ_2.

As shown in Fig. 28.12, the normalized Jones matrix of the incident ray can be expressed as

$$E_{in} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{28.16}$$

The polarization direction of the incident ray changes according to the polarizer and its polarization direction, which can be represented by rotating matrix R (φ):

$$R(\varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}. \tag{28.17}$$

The Jones-matrix vector after the incident ray passes through the polarizer is expressed as

$$E_1 = R(\theta_1)E_{in} = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 \\ -\sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin\theta_1 \\ \cos\theta_1 \end{bmatrix}. \tag{28.18}$$

Because the optical axis of the LC-SLM is parallel to the x-axis, its Jones matrix can be expressed as

$$SLM(\beta) = \begin{bmatrix} 1 & 0 \\ 0 & E^{-i\beta} \end{bmatrix}. \tag{28.19}$$

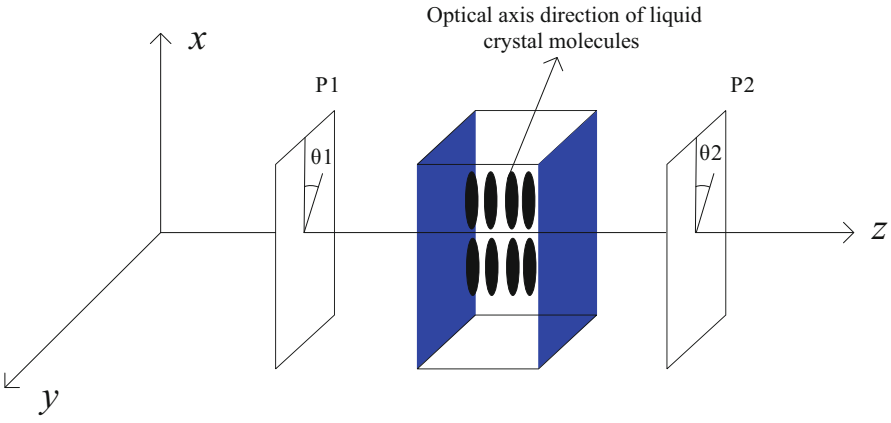


Fig. 28.12 LC-SLM working schematic diagram

The Jones matrix of the analyzer is

$$P = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{28.20}$$

The Jones matrix of the outgoing light after passing the analyzer is

$$E_2 = PR(-\varphi_2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\varphi_2 & \sin\varphi_2 \\ \sin\varphi_2 & \cos\varphi_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \sin\varphi_2 & \cos\varphi_2 \end{bmatrix}. \tag{28.21}$$

From the above derivation, it can be observed that after passing through the polarizer and LC-SLM, the Jones vector of the incident ray can be expressed as

$$\vec{E}_{\text{out}} = PR(-\varphi_2)SLM(\beta)R(\varphi_1)E_{\text{in}}. \tag{28.22}$$

Substituting Eqs. 28.12, 28.13, and 28.15 into the above formula, the complex amplitude of the outgoing light can be obtained as

$$\vec{E}_{\text{out}} = \sin\varphi_1 \sin\varphi_2 + \cos\varphi_1 \cos\kappa_2 e^{-j\beta}. \tag{28.23}$$

Simplifying Eq. 28.17, the light-intensity expression of the outgoing light is

$$I = \frac{1}{4}[\cos(\varphi_1 + \varphi_2) - \cos(\varphi_2 - \varphi_1) + \cos\beta(\cos(\varphi_1 + \varphi_2) - \cos(\varphi_2 - \varphi_1))]^2 + \frac{1}{4}\sin^2\beta[\cos(\varphi_1 + \varphi_2) - \cos(\varphi_2 - \varphi_1)]^2. \tag{28.24}$$

The phase-delay expression of the outgoing light is [96, 97]

$$\delta = \beta + \arctan \frac{\sin\beta[\cos(\varphi_1 + \varphi_2) + \cos(\varphi_2 - \varphi_1)]}{\cos(\varphi_1 + \varphi_2) - \cos(\varphi_2 - \varphi_1) + \cos\beta[\cos(\varphi_1 + \varphi_2) - \cos(\varphi_2 - \varphi_1)]}. \tag{28.25}$$

Application of Liquid-Crystal Spatial Light Modulators to Adaptive Optics

Phase-Screen Simulation of a Liquid-Crystal Spatial Light Modulator

A phase screen is a gray-release value matrix that simulates the phase change of atmospheric turbulence. This matrix is generally obtained from the phase function of the turbulence model that must be loaded. Taking a 256 × 256 LCD screen as an example, the LCD is fitted to a 256 × 256 lattice; then, each gray-release value on

the phase screen represents the phase value of the corresponding point on the lattice. The generated phase-screen matrix must consider the size of the LCD area and its correspondence with the analog phase-screen data [60].

The turbulence-effect simulation system is composed of a phase-type electrically addressable liquid-crystal spatial light modulator to simulate the far-field turbulence effect of laser signals [58]. The light emitted by the laser continuously attenuates the incident laser energy by controlling the relative angle of the polarizer. The wave plate and beam of the light-adjustment system are used to adjust the phase and spot size. By combining the light-beam attenuation, the light-spot size and laser power of the spatial light modulator are within the linear operating range of the system (Fig. 28.13).

The control computer generates a series of phase data as needed and loads these phase data on the liquid-crystal spatial light modulator. The phase-modulated laser is reflected by the mirror on the sampling CCD for recording, and the modulated laser light spot is obtained by simulating the turbulence effect. Temporal and spatial changes in energy distribution are investigated. The settings of the computer are controlled using the voltage of the liquid-crystal spatial light modulator in a lookup table. The relationship between the control voltage and the control pixel phase is used to adjust the spatial phase of the light beam reaching each pixel of the mirror to simulate the turbulence [60].

LC-SLM Coherent Optical Communication Wavefront-Correction Technology

In 2019, Ke Xizheng et al. from Xi'an University of Technology used a liquid-crystal spatial light modulator to correct the optical-signal distortion of a wave transmitted through atmospheric turbulence, thereby improving the performance of

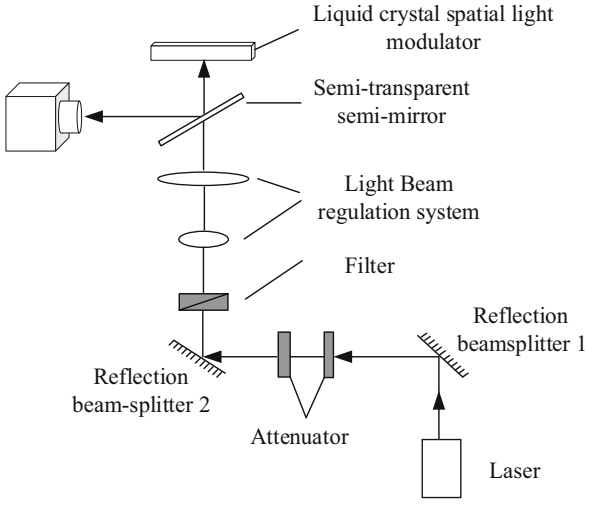


Fig. 28.13 Block diagram of the turbulence-effect simulation system

spatial coherent optical communication. The calibration experiment was divided into static wavefront-correction and field experiments [61].

Static Wavefront-Correction Experiment

The static wavefront-correction experiment adopts the conjugate-wavefront method. As shown in Fig. 28.14, first, the conjugate operation needs to be performed before the 1550-nm distorted wave. Then, the conjugate wavefront is calculated and processed, converted into the required gray-release value, and loaded into the liquid-crystal spatial light modulator to correct the distorted wavefront of the signal light.

The conjugate-wavefront method was used to correct the difference-frequency signal. To reduce the interference of external environmental factors on the difference-frequency signal, a wavefront-correction experiment was conducted at approximately 13 points. The correction results are shown in Fig. 28.15.

It can be seen from the figure that although the amplitude of the corrected difference-frequency signal fluctuates to a certain extent, in general, the amplitude tends to stabilize, and the fluctuation value of the voltage is not large over time; thus, there is not much difference between the amplitudes of the output sine wave. Moreover, the amplitude of the sine wave stabilizes at approximately 46 mV. Compared with the voltage amplitude before correction, the voltage amplitude after correction increased by approximately 1.7 times, and the signal-to-noise ratio after correction increased by nearly three times.

Figure 28.16 shows the output-voltage value of the difference-frequency signal before wavefront correction, where a is the voltage amplitude in the in-phase direction, and b is the voltage amplitude in the quadrature direction (Fig. 28.17).

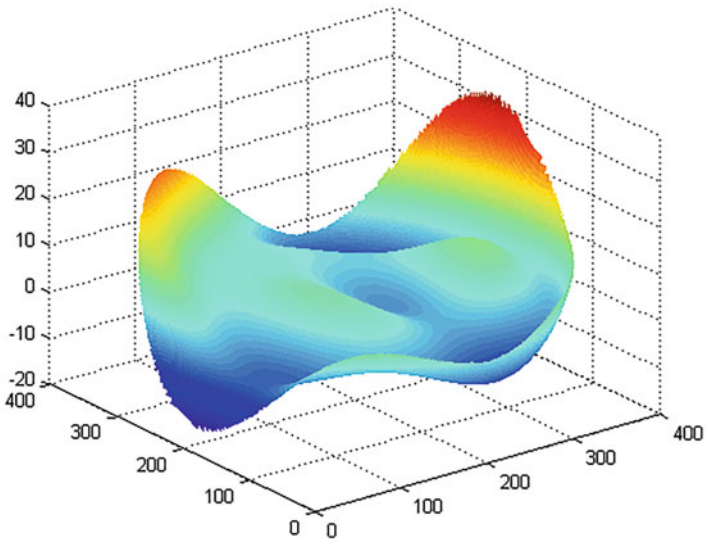


Fig. 28.14 1550-nm light-beam distorted-wave conjugate wavefront

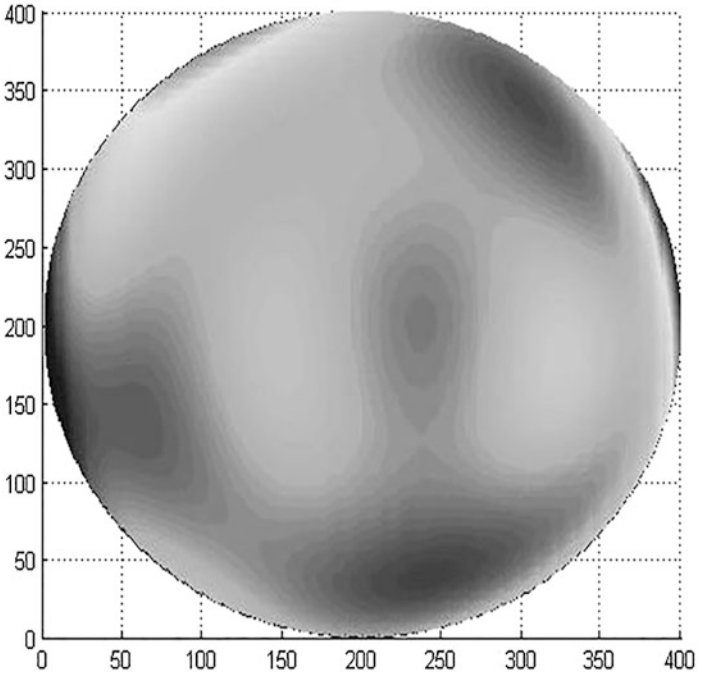


Fig. 28.15 Conjugate-wavefront gray-release diagram

1.3-Km Field Experiment

To further verify the influence of atmospheric turbulence on the signal-light wavefront and heterodyne signal, the laboratory conducted a space optical communication field experiment; the straight-line distance was 1.3 km.

In the experiment, the signal and beacon lights are coupled into the same optical fiber through the coupler and then transmitted by the optical antenna of the transmitting end. The receiving end receives the two beams of light and the sum of the beams of light, while a Shack–Hartmann wavefront sensor measures the beacon light. The wavefront of the beacon light fits the distorted wavefront of the beacon light in real-time.

According to the Zernike polynomial coefficients given in the fitting process, the conjugate wave of the distorted wavefront is obtained by the conjugate method. According to the modulation characteristics of the transmitted liquid-crystal spatial light modulator for the signal and beacon lights, the conjugate wavefront is converted into the gray-release value of the liquid-crystal spatial light modulator. It is then loaded onto the liquid-crystal spatial light modulator to correct the distorted wavefront of the signal light. Figures 28.18, 28.19, 28.20, 28.21, and 28.22 show the results of the field experiment from 12:00 am to 4:00 am on May 7, 2028.

Figure 28.18 shows the amplitude of the difference-frequency signal before and after correction. In the figure, the dotted sine wave represents the difference-

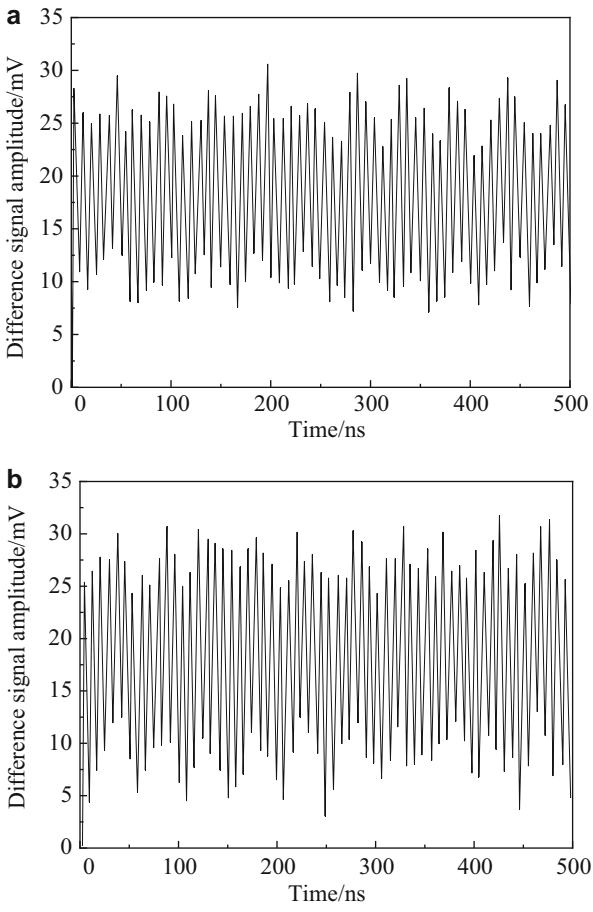


Fig. 28.16 Amplitudes of difference-frequency signal before correction. (**a**) In-phase; (**b**) quadrature [61]

frequency signal in the same phase direction, and the solid sine wave represents the difference-frequency signal in the quadrature direction. It can be observed from the figure that the amplitudes of the two difference-frequency signals output after correction are approximately twice as high as that before correction, and the amplitude of each sine signal tends to be constant.

Figure 28.19 shows the power-spectrum curve of the difference-frequency signal before and after correction. It can be seen from the figure that the main spectral peak of the power spectrum of the difference-frequency signal after correction becomes more acute. The peak value of the power-spectral density of the difference-frequency signal at the main spectral peak increases from 15.7 dB/Hz to 21.3 dB/Hz. The noise of the difference-frequency signal before correction is filtered out, and the system signal-to-noise ratio is improved.

As shown in Fig. 28.20, the amplitude of the difference-frequency signal of the system before correction is too small, and it fluctuates significantly in the adjacent

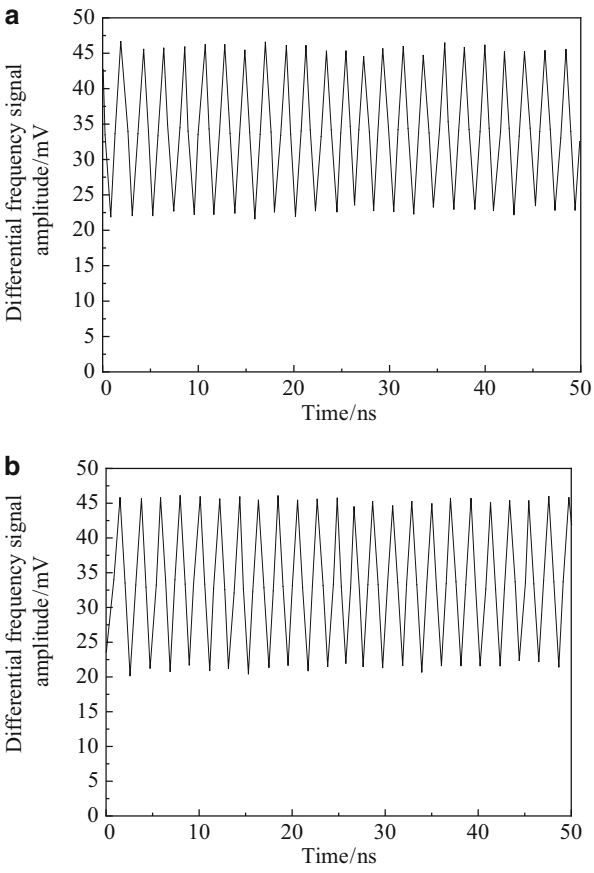


Fig. 28.17 Amplitude of the corrected difference-frequency signal. (**a**) in the same phase direction; (**b**) in the orthogonal direction [61]

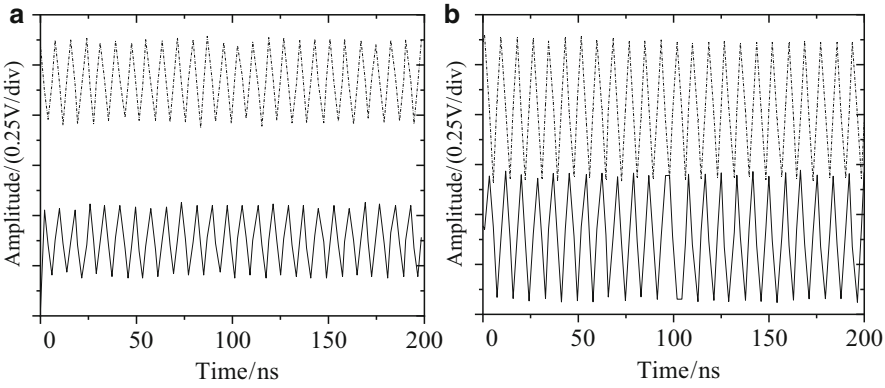


Fig. 28.18 Difference-frequency signal amplitude before and after correction [61]. (**a**) Before correction; (**b**) after correction

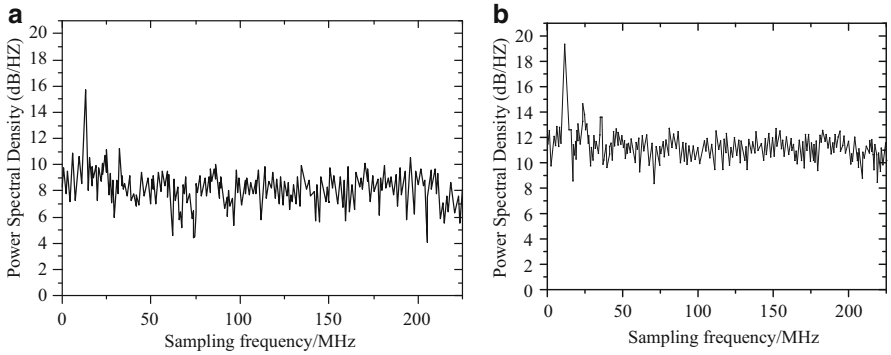


Fig. 28.19 Power spectrum of difference-frequency signal before and after correction [61]. (**a**) Before correction; (**b**) after correction

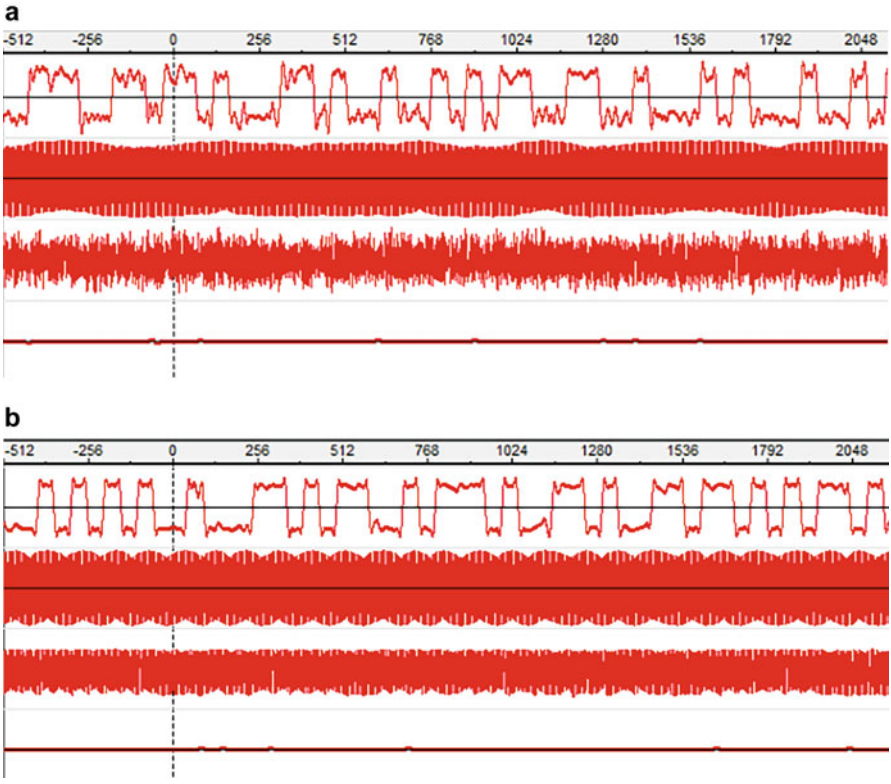


Fig. 28.20 Baseband signal before and after AO correction [61]. (**a**) Before correction; (**b**) after correction

period of the difference-frequency signal. Thus, it cannot meet the minimum amplitude requirements for decoding. The baseband signal output by the decoder is irregular in shape and has many surrounding glitches. After the signal optical-

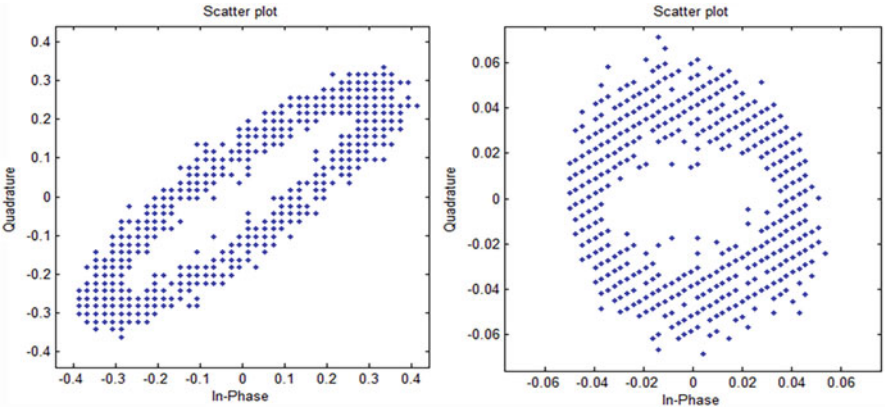


Fig. 28.21 Received signal constellation before calibration [61]

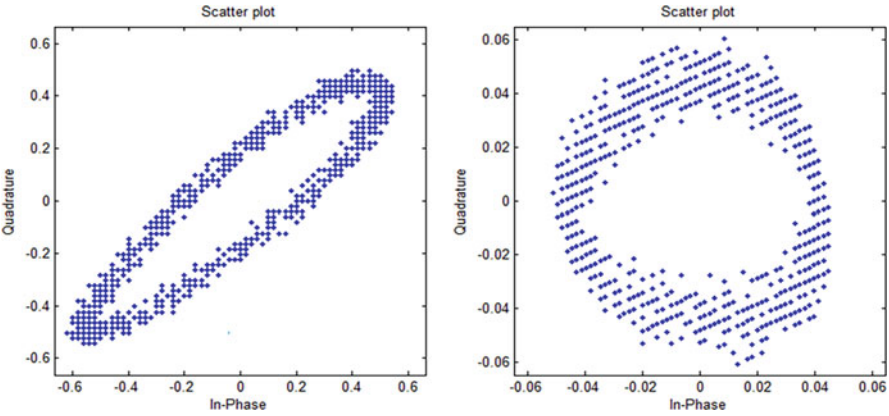


Fig. 28.22 Received signal constellation after correction [61]

wavefront distortion is corrected using AO technology, the amplitude of the difference-frequency signal meets the decoding requirements of the receiving end. The shape of the baseband signal decoded by the decoder and the square-wave approximation can be used to decode the transmitted video signal.

Comparing the constellation maps of the received signal before and after correction in Figs. 28.21 and 28.22, the dispersion between the points in the scatter map before correction is relatively large. After the adaptive-optics system corrects the distorted wave, the distribution of each scatter point in the constellation map of the received signal is relatively concentrated. The bit-error rate of the system after correction is reduced by approximately ten times.

The experimental results show that the amplitude of the photocurrent of the corrected differential-frequency signal can be increased by approximately 1.7 times, and the output signal-to-noise ratio can be increased by approximately three times using the transmissive liquid-crystal spatial light modulator for static

wavefront correction. The field-experiment results show that the amplitude of the output difference-frequency signal after correction increases once, and the bit-error rate of the demodulated signal is greatly reduced, which is approximately ten times less than that before correction.

Adaptive-Optical System of a Dual Liquid-Crystal Spatial Light Modulator

To solve the problem of energy loss of a single liquid-crystal spatial light modulator in the correction of an adaptive optics system, in 2022, Ke and Yang of Xi'an University of Technology designed an adaptive-optics system based on the structure of a dual liquid-crystal spatial light modulator. The dual liquid-crystal system compensates for the polarization-modulation loss caused by the liquid crystal.

The polarization-modulation characteristics of the liquid-crystal wavefront corrector cannot be avoided; however, dual liquid-crystal wavefront correctors can be used to modulate the light waves in the two polarization directions, as shown in Fig. 28.23. This can compensate for the energy loss caused by the polarization modulation of the liquid crystal. Notably, the beam of light combined after correction by the double corrector requires a higher optical-path modulation than that of the single corrector.

Figure 28.24 shows a schematic diagram of the experimental correction structure of the dual liquid-crystal spatial light modulator (LC-SLM) designed by Xi'an University of Technology. LC-SLM 1 distorts the Gaussian beam emitted by the light source. The beam-splitting prism divides the distorted light beam into two beams, one of which is corrected by LC-SLM 2, and the other is corrected by LC-SLM 3. The distorted light is captured by the CCD camera, and the gray-release diagram of the conjugate phase calculated by the Gerchberg–Saxton (GS) algorithm is used to load the light into LC-SLM 2 and LC-SLM 3, thereby realizing a closed-loop correction before the distorted wave [97].

Figures 28.25, 28.6, and 28.27 show light-spot diagrams before and after distorted light-spot correction for Gaussian light beams with strong, medium-strong, and weak atmospheric turbulence, respectively, using the GS algorithm.

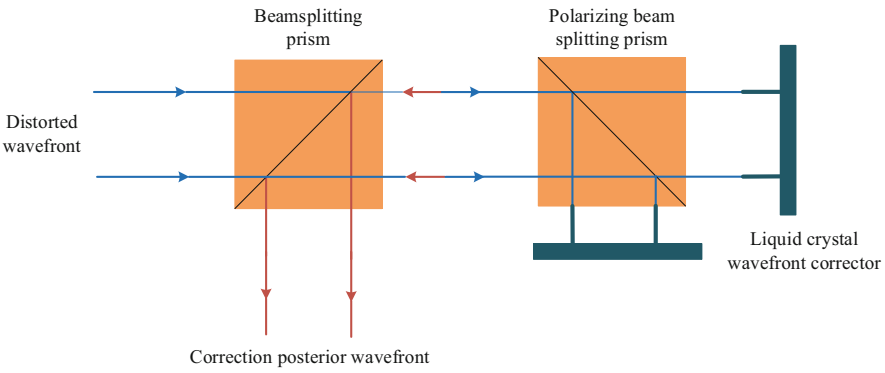


Fig. 28.23 Dual liquid-crystal wavefront corrector correction

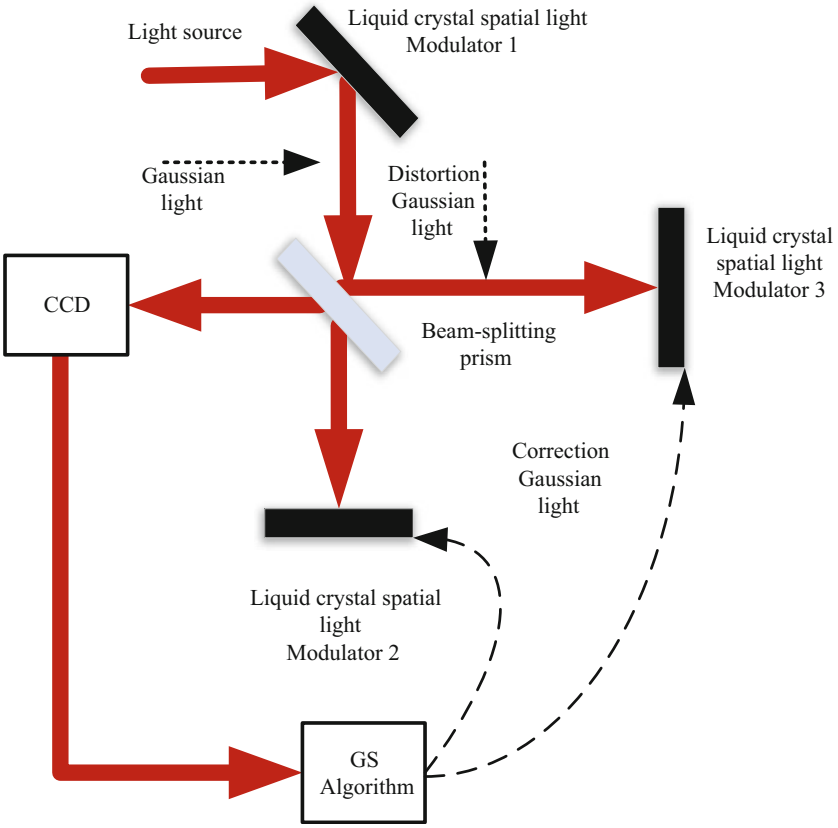


Fig. 28.24 Schematic diagram of the experimental correction structure of a dual liquid-crystal spatial light modulator for distorted light

It can be seen from Figs. 28.25, 28.26, and 28.27 that the Gaussian light beam becomes dispersed and broken after passing through the transmission environment under different turbulence intensities. However, after correction by the system, the light spot condenses, and the energy is greater. The convergence and quality are improved to a great extent.

Figure 28.28 shows the Strehl ratio (SR) variation curve of the distorted wavefront of the system under strong, medium-strong, and weak turbulence environments during simulation after 300 corrective iterations.

It can be seen from Fig. 28.28 that in a strong turbulent environment, the SR ratio of the system increases from 0.22 to 0.75; in a medium-strong turbulent environment, the SR of the system increases from 0.43 to 0.89; in a weak turbulent environment, the SR of the system increases from 0.73 to 0.97. The corrected SR value is improved to a certain extent, which shows that the correction of the dual liquid-crystal system has a certain effect.

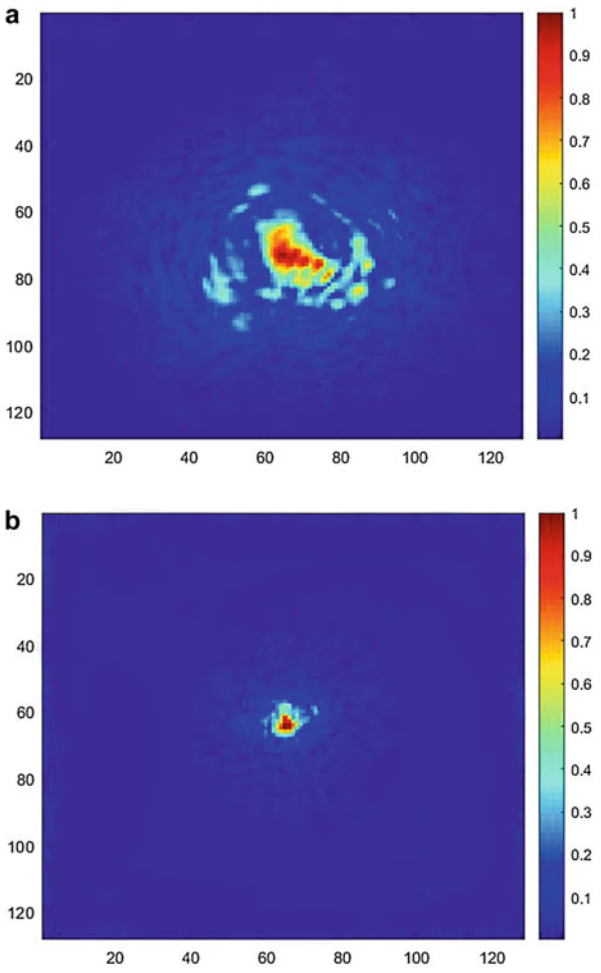


Fig. 28.25 Light-spot diagram before and after algorithm iteration when $D/r_0 = 20$. (**a**) Before iteration; (**b**) after iteration

Combined Liquid-Crystal–Deformable-Mirror Adaptive-Optics System

The deformable mirror has no dispersion problems and can meet the working requirements of a wide band; however, owing to the limitations of the manufacturing process and other reasons, the number of driving units is limited. This cannot meet the correction and imaging requirements of a large-aperture telescope system, and is not suitable for working in the visible and short-wave bands.

Because of its high-density pixel unit, the liquid-crystal spatial light modulator can meet the needs of a large-aperture telescope system for imaging and correction; however, because of its other characteristics, it is difficult to meet the correction and imaging requirements in the long-wave band. Therefore, in 2022, Ke and Yang of Xi'an University of Technology combined the advantages of deformable mirrors and

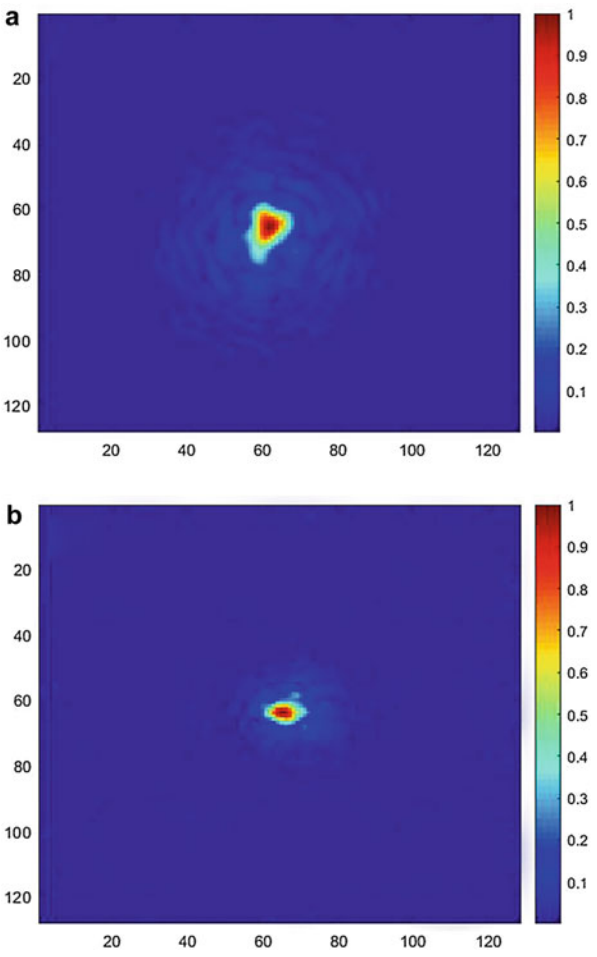


Fig. 28.26 Light-spot diagram before and after algorithm iteration when $D/r_0 = 10$. (**a**) Before iteration; (**b**) after iteration

liquid-crystal spatial light modulators and designed a set of adaptive systems based on both.

Figure 28.29 shows the experimental optical path of the deformable mirror–liquid-crystal spatial light modulator structure for correcting distorted light. LC-SLM 1 was used to distort the phase of the Gaussian light beam emitted by the light source, and the distorted light passed through an 800-nm dichroic chip. For the distorted light beam with a wavelength greater than 800 nm, the deformable-mirror correction loop used the eigenmode method for closed-loop correction. For the distorted light beam with a wavelength less than 800 nm, the liquid-crystal spatial light modulator loop used the GS algorithm for closed-loop correction [97].

Figures 28.30 and 28.31 show the light spots before and after the simulation of the system using the short-band correction loop. For the distorted light beam with

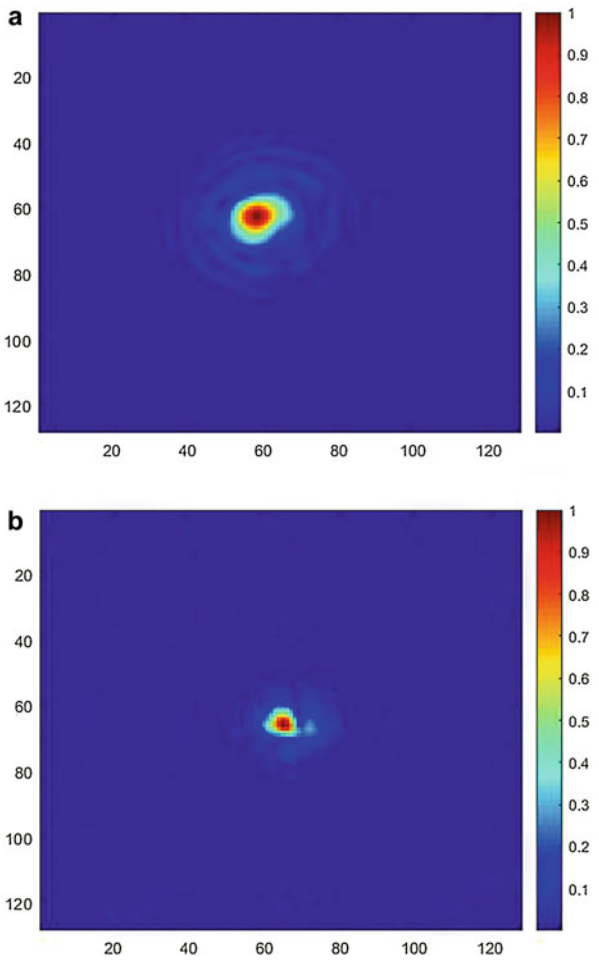


Fig. 28.27 Light-spot diagram before and after algorithm iteration when $D/r_0 = 2$. (**a**) Before iteration; (**b**) after iteration

wavelengths of 520 nm and 632.8 nm in a medium-strong turbulent environment, the GS algorithm iterates 200 times to make the correction. The change trend of the system SR can be observed.

Figures 28.32 and 28.33 show that the system adopts a long-band correction loop during the simulation. For the distorted light beam with wavelengths of 980 nm and 1550 nm in a medium-strong turbulent environment, the deformable-mirror eigenmode method iterates 20 times to make the correction. The light-spot diagram and change-trend diagram of the SR are shown.

From Figs. 28.30, 28.31, 28.32, and 28.33, it can be seen that for medium-strong turbulent environments, the distorted wavefronts of the long and short wavelengths are respectively corrected by the deformable-mirror correction loop and the liquid-crystal spatial light modulator correction loop. A good correction effect was

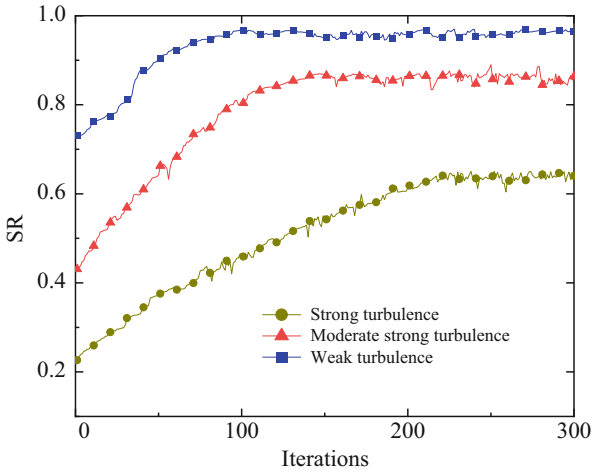


Fig. 28.28 SR variation curve of the system under different turbulence conditions

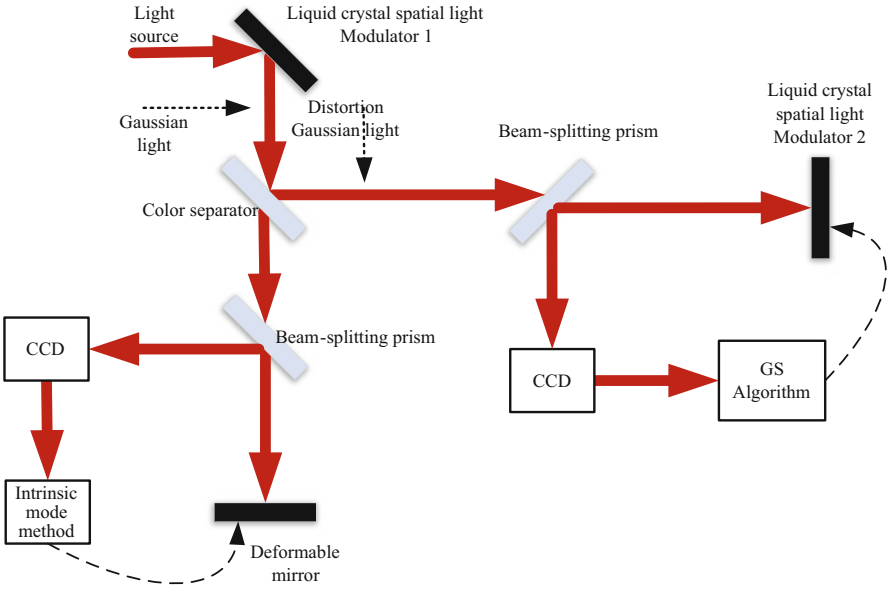


Fig. 28.29 Schematic diagram of the experimental setup for correcting distorted light with the deformable mirror–liquid-crystal spatial light modulator

obtained, and the light spot converged and was concentrated from its discrete and broken state.

In the medium-strong turbulent environment, the system SR value at the 520-nm wavelength increased from 0.37 to 0.74, that at the 632.8-nm wavelength increased from 0.38 to 0.79, that at the 980-nm wavelength increased from 0.45 to 0.75, and

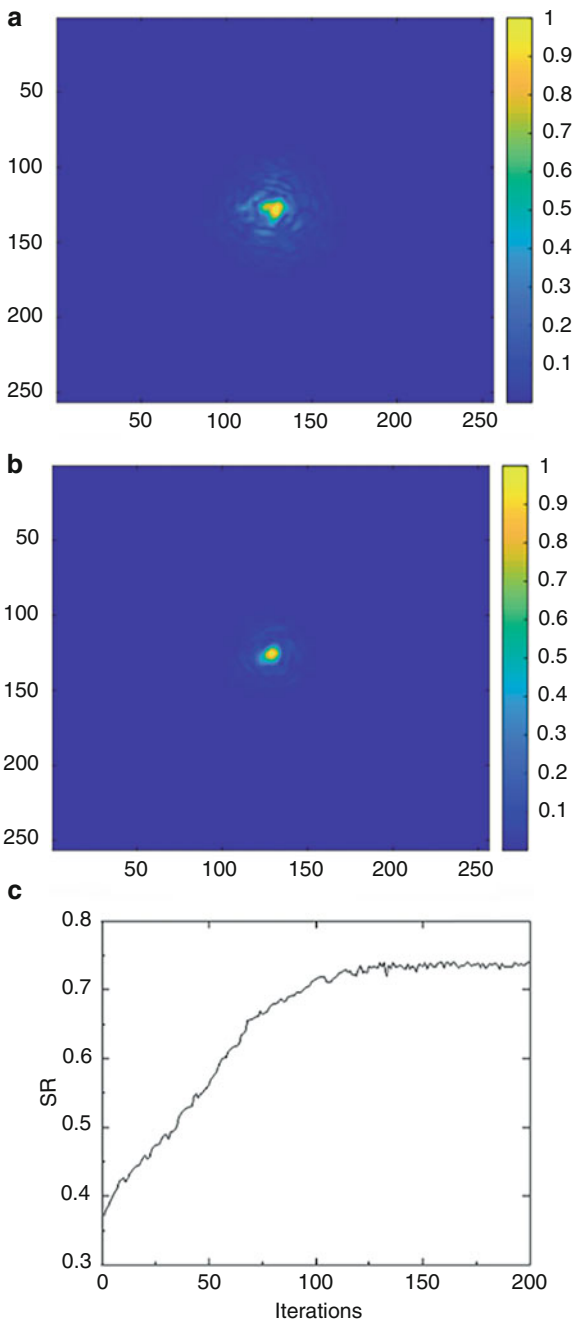


Fig. 28.30 Light spot and before and after algorithm iteration and SR curve when $l = 520$ nm. (**a**) Before correction; (**b**) after correction; (**c**) SR

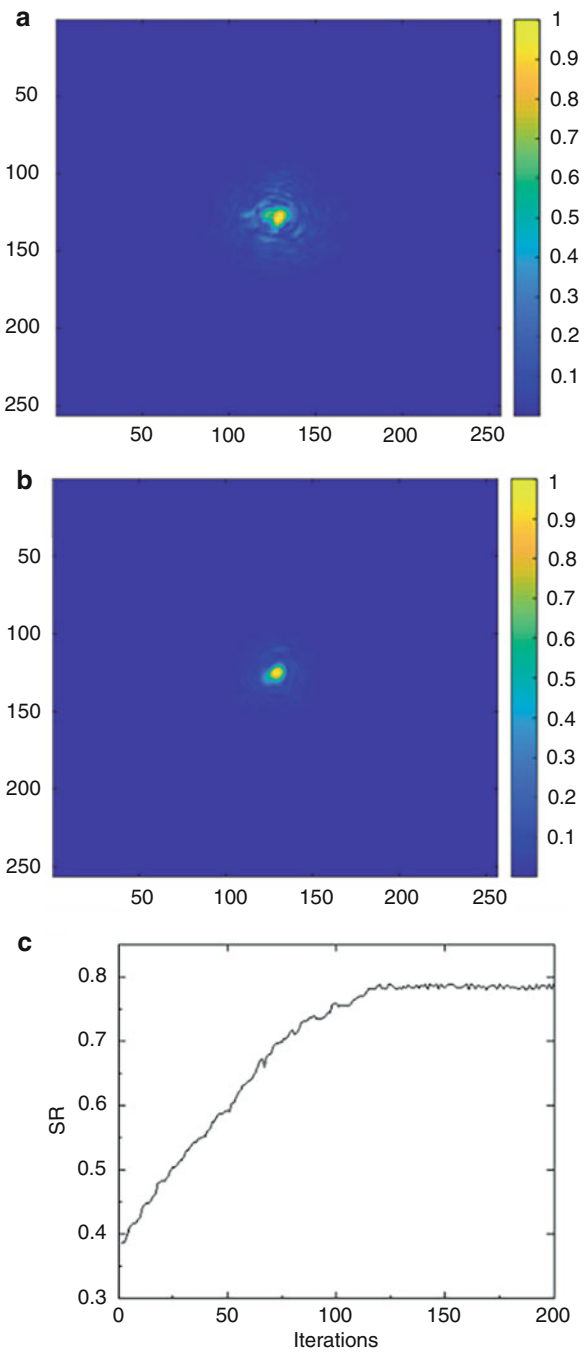


Fig. 28.31 Light spot before and after algorithm iteration and SR curve when $l =$ 632.8 nm. (**a**) Before correction; (**b**) after correction; (**c**) SR

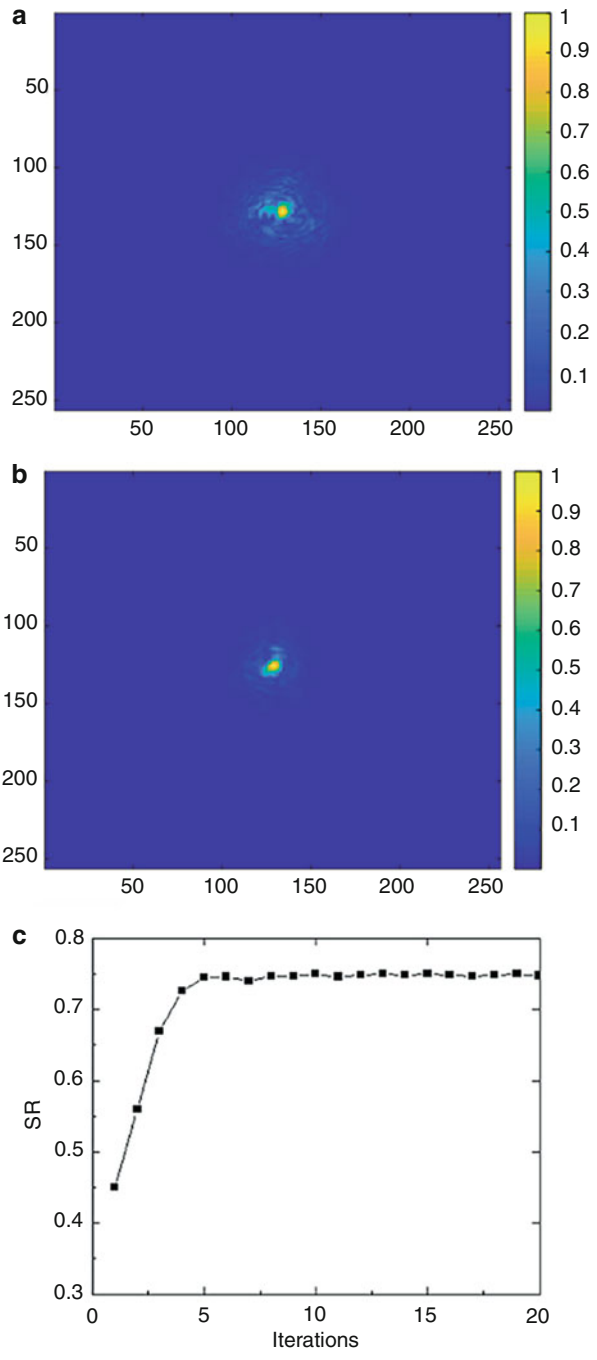


Fig. 28.32 Light spot and SR curve before and after algorithm iteration at $l =$ 980 nm. (**a**) Before correction; (**b**) after correction; (**c**) SR

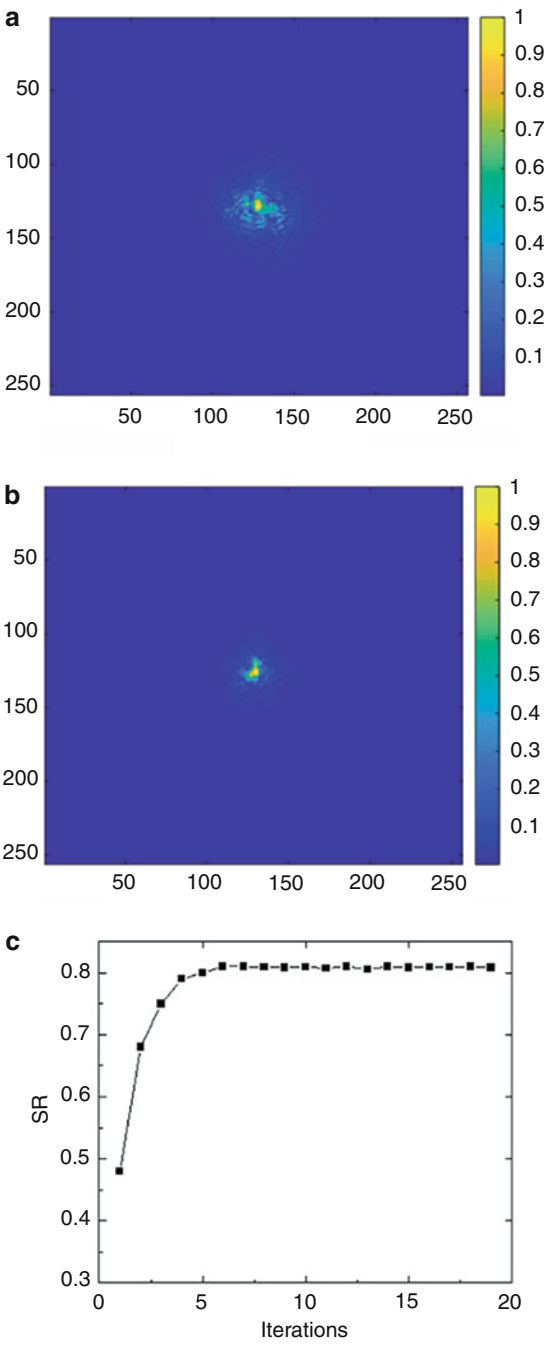


Fig. 28.33 Light spot and SR curve before and after algorithm iteration when l = 1550 nm. (**a**) Before correction; (**b**) after correction; (**c**) SR

that at the 1550-nm wavelength increased from 0.48 to 0.81. The simulation results showed that the post-SR value of each band was improved after the system was corrected. It also shows that the system helped to correct the distorted wavefront.

Summary

A liquid-crystal spatial light modulator is an ideal device for high-precision and high-resolution wavefront control; however, its applicability is low. Pure liquid-crystal spatial light phase modulators can achieve pure phase modulation greater than two without losing light intensity and have broad application prospects in the fields of static and dynamic wavefront control.

With the development of wavefront-control technology, the performance and accuracy requirements for wavefront controllers have increased. Therefore, the performance parameters of the spatial optical liquid-crystal phasers were comprehensively evaluated, and the errors in the devices were compensated and corrected.

High-precision wavefront correction is important. As an ideal device for high-precision and high-resolution wavefront control, liquid-crystal spatial light modulators have broad application prospects in static and dynamic wavefront control. To further optimize the performance of liquid-crystal spatial light modulators and improve the ability of adaptive optics to correct wavefront distortion, the following steps are necessary.

1. Under the action of an external electric field, a liquid-crystal spatial light modulator turns the liquid-crystal molecules to change the refractive index and modulates the wavefront. Because of the relaxation characteristics, the liquid-crystal molecules cannot be immediately steered; thus, their response time remains at the millisecond level. This affects the fast laser pointing and reduces the system performance of the liquid-crystal spatial light modulator. Therefore, methods should be studied to expedite the steering of the liquid-crystal molecules, improve the response speed of the liquid-crystal spatial light modulator, and improve the performance of the adaptive optical system.
2. Liquid-crystal spatial light modulators are used for wavefront correction. The phase-modulation accuracy of the spatial light modulator is affected by factors, such as the experimental environment. Therefore, the phase of the liquid-crystal spatial light modulator should be calibrated. The same liquid-crystal spatial light modulator can be calibrated multiple times using different methods to improve the measurement accuracy and wavefront correction.
3. When a liquid-crystal spatial light modulator was used in the dynamic wavefront-correction experiment, its response speed was slow, and the reading delay of the gray-release image was large. Therefore, to generate a more accurate gray-release phase image, a nonlinear phase-compensation method can be used, or the existing phase-recovery algorithm can be further improved to increase the calculation efficiency and optimize the homogenization effect.

References

1. Haung L H, Rao C H, Jiang W H, et al. Modified Gaussian influence function of deformable mirror actuators [J]. Optics Express, 2008, 16(1): 108–114.
2. Cheng Xin, Ren Xiuyun, Han Yujing, et al. Grating diffraction efficiency based on liquid crystal spatial light modulator [J]. Chinese Journal of Photonics, 2006, 35(4): 603–607.
3. Ma Sijin. Testing and application research on phase modulation characteristics of liquid crystal spatial light modulator [D]. Beijing: Beijing University of Technology, 2014.
4. Love Gordon D. Wave-front correction and production of Zernike modes with a liquid-crystal spatial light modulator [J]. Applied Optics, 1997, 36(7): 1517–1528.
5. Gu D F, Winker B, Wen B, et al. Wavefront control with a spatial light modulator containing dual-frequency liquid crystal [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2004, 5553(12): 68–82
6. Olivier S S, Kartz M W, Bauman B J, et al. High-resolution wavefront control using liquid crystal spatial light modulators [J]. Proceedings of SPIE - The International Society for Optical Engineering, 1999, 3760(20): 47–51
7. Zhang Shuyan. Research on characteristics and applications of ferroelectric liquid crystal devices [D]. Shanghai: Fudan University, 2000.
8. Matkin W B. Steered agile beams program support for Army requirements [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2002, 4489(4): 1–12.
9. Mcmanamon P. An overview of optical phased array technology and status [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2005, 5947(13): 152–161.
10. Sharp R C, Resler D P, Hobbs D S, et al. Electrically tunable liquid-crystal wave plate in the infrared [J]. Optics Letters, 1990, 15(1): 87–89.
11. Lin Y H, Voelz D G, Ricklin J C, et al. Compact 4 cm aperture transmissive liquid crystal optical phased array for free-space optical communications [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2005, 5892(20): 107–118.
12. Linnenberger A, Serati S, Stockley J, et al. Advances in optical phased array technology [J]. Proceedings of Spie the International Society for Optical Engineering, 2006, 6304(7): 199–207.
13. Savage N. Digital spatial light modulators [J]. Nature Photonics, 2009, 3(3): 170–172.
14. Engström D, O'Callaghan M J, Walker C, et al. Fast beam steering with a ferroelectric-liquid-crystal optical phased array [J]. Applied Optics, 2009, 48(9): 1721–1730.
15. Huang Daquan, Xie Zhixing, Liu Yuxiang. Projection system of amorphous silicon liquid crystal light valve [J]. Vacuum Electronic Technology, 1994, 3 (1): 1–6.
16. Kangcheng Q, Jianbo C, Changjun G, et al. A-Si:H-based LCLV with an CdTe and VOPc multiple absorption layers [J]. IEEE Transactions on Electron Devices, 2004, 51(5):823–829.
17. Guo Huaiwen. Amplitude modulation characteristics and applications of liquid crystal spatial light modulators [D]. Beijing: Chinese Academy of Engineering Physics, 2023.
18. Fan Jiaxin. Research on beam of light pointing control method of liquid crystal spatial light modulator [D]. Changchun: Changchun University of Science and Technology, 2024.
19. Zhang Jian, Xu Lin, Wu Liying, et al. Deflection of programmable beams of light in liquid crystal optical phased array [J]. Chinese Journal of Photonics, 2008, 37 (8): 1497–1502.
20. Huang D, Fan W, Li X, et al. An optically addressed liquid crystal light valve with high transmittance [C]. Holography, Diffractive Optics, and Applications V. China: SPIE, 2012: 209–215.
21. Huang Dajie, Fan Wei, Zhang Panzheng, et al. Development progress and performance test of optically addressable spatial light modulators [J]. Infrared and Laser Engineering, 2016, 45 (9): 230–235.
22. Wu L, Wang X, He X, et al. Arbitrary multiple beam forming by two cascaded liquid crystal optical phased arrays [J]. Optics Express, 2020, 26(13): 17066–17077.
23. Li S Q, Xu X, Maruthiyodan Veetil R, et al. Phase-only transmissive spatial light modulator based on tunable dielectric metasurface [J]. Science, 2019, 364(6445): 1087–1090.

24. Dai Haitao. Research on the characteristics and applications of LCOS phase space light modulator [D]. Shanghai: Fudan University, 2005.
25. Niu Qifeng, Wang Chunyang, Zhang Guangping, et al. Solving the vector distribution of liquid crystal phased array and its beam deflection [J]. Optical Precision Engineering, 2020, 26 (12): 2894–2901.
26. Wang Ning, Li Guohua, Yun Maojin, et al. Study on the birefringence characteristics of liquid crystal electronically controlled [J]. China Laser, 2002, 29 (12): 1064–1066.
27. Liu Jing. Twisted nematic liquid crystal pure phase spatial light modulator [D]. Dalian: Dalian University of Technology, 2008. Liu Jing. Twisted nematic liquid crystal pure phase spatial light modulator [D]. Dalian: Dalian University of Technology, 2008.
28. Liu Yongjun, Xuan Li, Hu Lifa, et al. Research on high-precision pure-phase liquid crystal spatial light modulator [J]. Chinese Journal of Optics, 2005, 25 (12): 1682–1686.
29. Lu Qiang, Sheng Lei, Zhang Xin, Bai Jiqing, et al. Research on pure phase modulation characteristics of liquid crystal spatial light modulators under oblique incidence [J]. China Laser, 2016, 43 (01): 190–196.
30. Xie Yuzhang, Liquid Crystal Physics [M]. Beijing: Science Press, 1988.
31. Krueger S, Kamps J, Wernicke G K G, et al. Spatial light modulator system for application as a dynamic diffractive element and in optical image processing [C]. Diffractive/Holographic Technologies and Spatial Light Modulators VII. Germany: SPIE,2000: 179–190.
32. Armitage D, Thackara J I, Eades W D, et al. Photoaddressed liquid crystal spatial light modulators [J]. Applied Optics, 1989, 28(22): 4763–4771.
33. Broomfield S E, Neil M A A, Paige E G S, et al. Four-level, phase only, spatial light modulator [J]. Electronics Letters, 1993, 20(29): 1661–1663.
34. Yan Sai, Wu Xiu, Wang Chunyun, et al. Experimental study on amplitude and phase modulation characteristics of spatial light modulators [J]. Chinese Journal of Quantum Optics, 2017, 24 (1): 107–112.
35. Liu Dongmei, Liu Juan, Zhang Zhao, et al. Calibration of the initial phase and multi-wavelength response of phase-type liquid crystal spatial light modulators [J]. Advances in Laser and Optoelectronics, 2012, 49 (4): 84–87.
36. Chen Huaixin, Sui Zhan, Chen Zhenpei, et al. Optical modulation characteristics of LCD TV LCTV and its application [J]. China Laser, 2000, 27 (8): 741–745.
37. Li Jie, Tang Feng, Wang Xiangchao, et al. Grating transverse shear interferometer and its systematic error analysis [J]. China Laser, 2014, 41 (5): 201–210.
38. Yamauchi M, Marquez A, Davis J A, et al. Interferometric phase measurements for polarization eigenvectors in twisted nematic liquid crystal spatial light modulators [J]. Optics Communications, 2000, 201(3): 1–6.
39. Wang Yijia, Zhang Chonglei, Wang Rong, et al. Phase-difference surface plasmon resonance sensor based on Mach-Zehnder structure [J]. China Laser, 2013, 40 (7): 215–221.
40. Ge Aiming, Sui Zhan, Xu Keyang, et al. Study on pure phase modulation characteristics of reflective LCOS devices [J]. Chinese Journal of Physics, 2003, 52 (10): 2481–2485.
41. Zhang H, Zhang J, Wu L. et al. Evaluation of phase-only liquid crystal spatial light modulator for phase modulation performance using a Twyman–Green interferometer [J]. Measurement Science and Technology, 2007, 20(6): 1724–1733.
42. Liu Hua, Lu Zhenwu, Li Fengyou, et al. Large-aperture aspheric computational hologram detection system [J]. Infrared and Laser Engineering, 2006, 35 (2): 177–202.
43. Burge J H, Anderson D S. Full-aperture interferometric test of convex secondary mirrors using holographic test plates [C]. Advanced Technology Optical Telescopes V. United States: SPIE,1994: 201–192.
44. Zhang Hongxin. Research on characteristic test method and wavefront correction of phase-type liquid crystal spatial light modulator [D]. Harbin: Harbin Institute of Technology, 2009.
45. Cao Z, Xuan L, Hu L, et al. Investigation of optical testing with a phase-only liquid crystal spatial light modulator [J]. Optics Express, 2005, 13(4): 1059–1065.

46. Zhao Xiaofeng, Li Dahai, Chen Zhenpei, et al. Measurement of phase modulation characteristics of liquid crystal spatial light modulators by radial shear interferometry [J]. Journal of Sichuan University (Natural Science Edition), 2002, 39 (4): 675–679.
47. Ren Xiuyun, Cai Chunwei, Wang Zhu, et al. Common measurement of phase modulation characteristics of LCSLM by transverse shear interferometry [J]. Chinese Journal of Photonics, 2007, 36 (5): 899–904.
48. Dou R, Giles M K. Closed-loop adaptive-optics system with a liquid-crystal television as a phase retarder [J]. Optics letters, 1995, 20(14): 1583–1585.
49. Lu Kanghua, Saleh B E A. Theory and design of the liquid crytal TV as an optical spatial phase modulator [J]. Optical Engineering. 1990, 29(3):240–246.
50. Wu Yunyun. Research on the application of adaptive optics technology in atmospheric optical communication [D]. Beijing: Chinese Academy of Sciences (Institute of Optoelectronic Technology), 2013.
51. Zhang Zhilu. Research on indoor atmospheric turbulence simulation system [D]. Taiyuan: Taiyuan University of Technology, 2019.
52. Rao Ruizhong. Modern Atmospheric Optics [M]. Beijing: Science Press, 2012.
53. Suo Qiangbo, Li Jinhong, Chen Linying, et al. Quality analysis of vortex beam of light in atmospheric turbulent propagation [J]. Optoelectronics. Laser, 2015, 26 (9): 2028–2034.
54. Zhang Wei, Ge Yaojun. Analysis of vortex shedding of bridge section based on modified turbulence freezing assumption [J]. Journal of Tongji University (Natural Science Edition), 2008, 36 (10): 1307–1313.
55. Li J, Yang S, Guo L, et al. Electromagnetic waves propagation in hypersonic turbulence using fractal phase screen method [J]. Journal of Electromagnetic Waves & Applications, 2017, 31(3): 250–262.
56. Harding C M, Johnston R A, Lane R G, et al. Fast simulation of a kolmogorov phase screen [J]. Applied Optics, 1999, 38(11): 2161–2170.
57. Bo Bin, Menk Nemle, Zhao Jianlin, et al. Generating vortex beams of light using a reflective pure-phase liquid crystal air-light modulator [J]. Optoelectronics. Laser, 2012, 23 (1): 74–78.
58. Li Hua, Chen Qianrong, Wang Yanbin, et al. Dynamic simulation of laser atmospheric turbulence effect using liquid crystal spatial light modulation [J]. Optoelectronics. Laser, 2016, 27 (7): 709–715.
59. Ren Zhenni. Atmospheric channel model research and system design for free space optical communication [D]. Guangzhou: Guangdong University of Technology, 2006.
60. Vorontsov M A, Carhart G W, Cohen M, et al. Adaptive optics based on analog parallel stochastic optimization: analysis and experimental demonstration [J]. Journal of the Optical Society of America A, 2000, 17(8): 1440–1453.
61. Kong Yingxiu. LC-SLM wavefront correction technology for spatial coherent optical communication [D]. Xi'an: Xi'an University of Technology, 2021.
62. Ke Xizheng, Han Kona. Wavefront simulation and wavefront correction of liquid crystal spatial light modulators [J]. Advances in Laser and Optoelectronics, 2019, 56 (05): 170–175.
63. L Wu,X Wang,C Xiong, et al. Polarization-independent two-dimensional beam steering using iquid crystal optical phased arrays [J].Chinese Optics Letters, 2017, 15(10): 85–88.
64. Zhao Dazun. Spatial light modulator [M]. Beijing: Beijing Institute of Technology Press. 1992.
65. Wang Jiangtao, Chen Xiangning. A-Si: H/FLC optical addressing liquid crystal light valve performance research [J]. Journal of Sichuan Ordnance Engineering, 2011, 32 (03): 127–129 + 146.
66. McManamon P F, Dorschner T A, Corkum D L, et al. Optical phased array technology [J]. Proceedings of the IEEE, 1996, 84(2): 268–298.
67. Lindquist R G, Kulick J H, Nordin G P, et al. High-resolution liquid-crystal phase grating formed by fringing fields from interdigitated electrodes [J]. Optics letters, 1994, 19(9): 670–672.
68. Titus C M, Bos P J. Efficient, polarization-independent, reflective liquid crystal phase grating [J]. Applied Physics Letters, 1997, 71(16): 2239–2241.

69. Nakamura H A. Model of Image Display in the optimized overdrive method for motion Picture quality improvements in liquid crystal devices [J]. Japanese Journal of Applied Physics, 2001,40(11):6435–6440.
70. Harris S R, Gonglewski J D, Vorontsov M A, et al. Numerical optimization of the performance of nematic liquid crystal optical phased arrays [J]. Proceedings of SPIE - The International Society for Optical Engineering, 2003, 5162(12): 157–171.
71. Wang X, Wang B, Bos P J, et al. Modeling and design of an optimized liquid-crystal optical phased array [J]. Journal of Applied Physics, 2005, 98(7): 1–8.
72. Barrick G F, Bos P J, Titus C M, et al. Computing the liquid crystal director field in optical phased arrays [J]. Optical Engineering, 2004, 43(4): 924–932.
73. Tholl H D. Novel laser beam steering techniques [C]. Technologies for Optical Countermeasures III. Germany: SPIE, 2006: 51–64.
74. Xun X, Cho D J, Cohn R W. et al. Spiking voltages for faster switching of nematic liquid-crystal light modulators [J]. Applied Optics, 2006, 45(13): 3136–3143.
75. Engstrom D, Bengtsson J, Eriksson E, et al. Improved beam steering accuracy of a single beam with a 1D phase-only spatial light modulator [J]. Optics express, 2008, 16(22): 20275–20287.
76. Kim J, Miskiewicz M N, Serati S, et al. Demonstration of large-angle nonmechanical laser beam steering based on LC polymer polarization gratings [C] Acquisition, Tracking, Pointing, and Laser Systems Technologies XXV. United States: SPIE, 2011: 210–224.
77. Love G D, Thalhammer G, Padgett M J, et al. Speeding up liquid crystal SLMs using overdrive with phase change reduction [J]. Optics Express, 2013, 21(2):1779–1797.
78. S Serati, C L Hoy, L Hosting, et al. Large-aperture, wide-angle nonmechanical beam steering using polarization gratings [J]. Optical Engineering, 2017,56(3):031211–031224.
79. X Xiang, J Kim, M.J. Escuti, et al. Bragg polarization gratings for wide angular bandwidth and high efficiency at steep deflection angles [J]. Scientific Reports, 2020, 8(1):467–475.
80. He A, Quan C. Wavefront correction for spatial nonuniformity of the liquid crystal on silicon based spatial light modulator [J]. Optics and Lasers in Engineering, 2019, 121(20): 377–388.
81. Kong Lingjang, Yi Wei, Yang Jianyu, et al. Scanning accuracy analysis of lidar liquid crystal phased array components [J]. China Laser, 2009, 36 (5): 1080–1085. (in Chinese)
82. Song Yan. Research on beam control model and algorithm of liquid crystal phased array [D]. Chengdu: University of Electronic Science and Technology of China, 2010.
83. Zou Yongchao. Research on deflection technology of large angle beam of light based on liquid crystal phased array [D]. Changsha: National University of Defense Technology, 2011.
84. Yin Xiaohu. Research on wave control data optimization algorithm of lidar liquid crystal phased array [D]. Chengdu: University of Electronic Science and Technology of China, 2011.
85. Zhou Yan, Kong Lingjiao, Pan Fei, et al. Pointing accuracy correction algorithm for liquid crystal phased array based on slope approximation [J]. China Laser, 2012, 39 (2): 1–6.
86. Qiao Tianyuan. Research on the accuracy of liquid crystal phased array imaging lidar [D]. Harbin: Harbin Institute of Technology, 2014.
87. Yang He, Tong Shoufeng, Zhang Lei, et al. Beam of light deflection control technology based on liquid crystal spatial light modulator [J]. Advances in Laser and Optoelectronics, 2017, 54 (11): 84–91.
88. Xiao F, Kong L. Angular magnification method of liquid crystal optical phased array based on telescope system [C]. International Symposium on Photoelectronic Detection and Imaging 2013: Laser Communication Technologies and Systems. China: SPIE, 2013: 629–636.
89. Li Song Vibration. Design of liquid crystal polarization grating and research on its light deflection characteristics [D]. Changchun: Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, 2021.
90. Niu Qifeng. Research on key technologies of light deflection of liquid crystal phased array [D]. Changchun: Changchun University of Science and Technology, 2023.
91. Cao Y, Chong L, Wu K H, et al. Dynamic coloration of polymerized cholesteric liquid crystal networks by infiltrating organic compounds [J]. Chinese Optics Letters, 2022, 20(9): 47–52.

92. Wu Xiaofang. Research on scintillation phenomenon of pure phase Holoeye pluto liquid crystal spatial light modulator [D]. Fuzhou, Fujian: Fujian Normal University, 2019.
93. Wang Dong. Research on agile control technology of laser beam based on pure-phase liquid crystal spatial light modulator [D]. Harbin: Harbin Institute of Technology, 2013.
94. Ignacio M, Jeffrey A. Davis, et al.Transmission and Phase Measurement for Polarization Eigenvectors in Twisted-nematic Liquid Crystal Spatial Light Modulators [J]. Optical Engineering, 1998, 37(11): 3048–3052.
95. Gibson J S, Chang C C, Ellerbroek B L. Adaptive Optics Wavefront Reconstruction by Adaptive Filtering and Control [C]. IEEE Conference on Decision and Control. United States: IEEE, 1999:761–766
96. Hilaire P S, Benton S A, Lucente M, et al. Synthetic aperture holography: a novel approach to three-dimensional displays [J]. Journal of the Optical Society of America A, 1992, 9(11): 1969–1977.
97. Wu J, Ke X, Yang Y, et al. Correction of Distorted Wavefront Using Dual Liquid Crystal Spatial Light Modulators [J] Photonics, 2022, 9(6): 426–437.

Mixers

29

Contents

Introduction 1066
Tube Mixers 1067
RF Mixers 1068
  Schottky-Diode Mixers 1068
  CMOS Mixers 1074
Mixer Circuit Structures 1075
  Single-Ended Structures 1075
  Single-Balanced Structures 1076
  Double-Balanced Structures 1077
  Triple-Balanced Structures 1079
  Mixer Linearity 1079
Optical Mixers 1081
  Working Principle 1081
  Research Progress and Current Status 1083
  3-dB Coupler and Phase Shifters 1083
  Mach–Zehnder Interference Optical Mixers 1084
  Polarized Optical Mixers 1087
  Crystal Optical Mixers 1090
  Other Types of Optical Mixers 1094
  Multimode Interference Coupler Optical Mixer 1095
  MMI Optical Mixer Classification 1097
  Arrayed Waveguide Grating Structured Optical Mixer 1099
Comparison of the Material Properties of Optoelectronic Devices 1101
  Common Materials for Optical Mixers 1101
  Comparison of Material Properties 1102
  Mixing-Error Analysis 1102
Analysis of the Mixing Efficiency of an Optical mixer in a Coherent Optical Communication System 1104
  Light Field Pattern Matching 1104
  Light-Spot Size Mismatch 1107
  Wavefront Mass 1108
  Vibration and Tracking Errors 1110
  Polarization Characteristics 1111

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_29

Research on Optical Mixers at Xi'an University of Technology 1113
Structural Design of an Optical mixer 1114
Structural Optimization 1115
Experimental Verification 1115
Summary 1116
References 1119

Abstract

Mixer is an indispensable component in electronic communication system, and it is the frontend circuit of heterodyne receiver and microwave measuring instrument, which is used in microwave communication, radar, and many microwave measurement systems. With the development of optical communication, optical mixing devices in coherent optical communication also play an important role in coherent receiving terminals and optical measurement systems.

Starting with the electric mixer, this paper describes the working principle and development status of the optical mixer and the terahertz mixer and briefly introduces the work of Xi'an University of Technology in the field of optical mixer. Finally, the future development trend and prospect of the mixer are prospected.

Keywords

Microwave circuit · Coherent optical communication · Microwave mixer · Optical mixer

Introduction

The world's first microwave mixer was developed by Armstrong in 1924, when it was mainly used to increase the detection range and reception sensitivity of radar [1]. Mixers at that time were frequency conversion elements based on variable resistance nonlinear diodes; however, the slow development of electron tube devices led to the slow development of mixers [2].

Mixers are mainly used in electronic countermeasures, radar, radio measurement, and coherent optical communication and can also be used as frequency multipliers, frequency dividers, and phase discriminators. As the key device in receiving end points, research on microwave mixers is of great significance in optical communication systems.

A mixer has a circuit structure that uses the nonlinear effect of nonlinear devices to transform signals of different frequencies or local oscillator signals into differential frequency signals, or sum frequency signals of the two [3]. Frequency conversion circuits can be divided into up converters and downconverters, according to the numerical relationship between the input signal, local oscillator signal, and output signal. The down converter is also called a mixer [4]. The mixer loads the received signal carrier and the local oscillator signal together onto a nonlinear device

to mix them. This converts the radiofrequency (RF) to an intermediate frequency signal. The nonlinear effect of the diode is essential to the mixing effect.

According to the principle of nonlinear devices, mixers can be divided into active devices and passive devices. A mixer using transistors and field effect transistors (FETs) as the infrastructure is called an active device mixer. It can obtain a higher gain and the combined interference is small; however, the circuit is complex and requires a DC bias.

A mixer using a diode as the infrastructure is called a passive device mixer or a resistive mixer. This type of electric mixer has a simple structure, easy integration, stable performance, and consistent electrical characteristics. Therefore, nonlinear electric mixers are mainly used. In addition, according to the circuit structure, electrical mixers can be divided into single-ended mixers, single-balanced mixers, double-balanced mixers, etc.

Tube Mixers

In 1918, Armstrong developed the superheterodyne architecture, which enabled an easy-to-use radio. The superheterodyne receiver had gain capability and additional noise reduction advantages and also provided an intermediate frequency current signal. It was the prototype of the early mixer concept based on the electron tube structure [5].

As shown in Fig. 29.1, the vacuum tube tunes the input signal and the independent heterodyne device (local oscillator signal) into a combined oscillation signal, and the multi-vacuum tube forms a high-frequency amplifier to heterodyne process and detects the combined signal. Through this method, the radiofrequency signal can be efficiently converted into a stable output intermediate frequency signal. The signal can be amplified simultaneously, which is the sum–difference product of a typical mixer. The circuit schemes of the electron tube mixer and the transistor mixer are compared, as shown in Fig. 29.2.

According to Armstrong's patent [6], the first-stage circuit structure consists of a vacuum tube rectifier system that utilizes the rectifying properties of the tube to

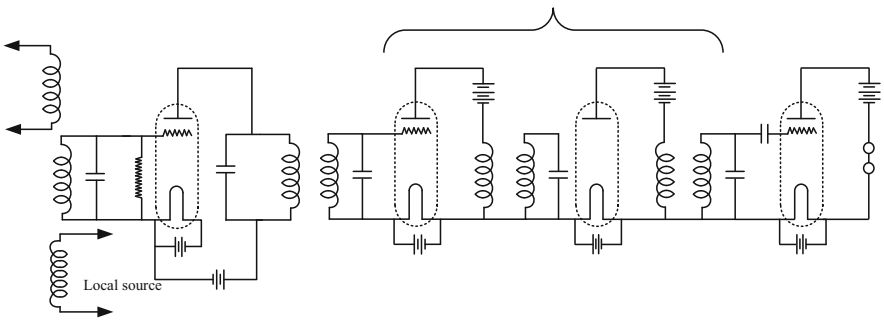


Fig. 29.1 Vacuum tube mixer structure [5]

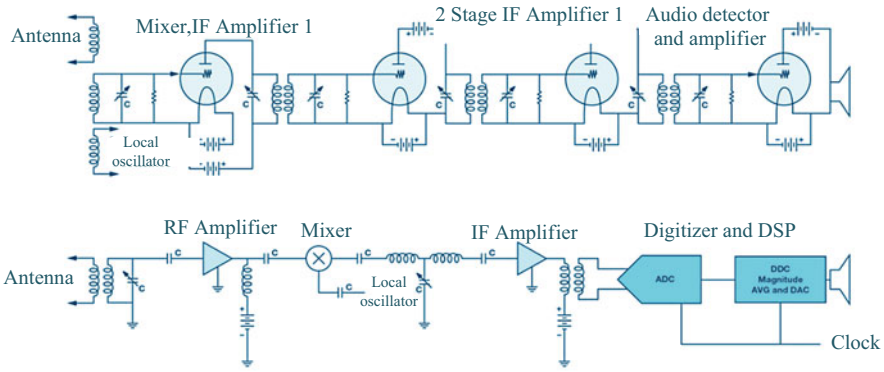


Fig. 29.2 Comparison of the tube mixer and transistor mixer [5]

generate a typical mixing product that mixes the target signal with the local oscillator signal. The only fundamental difference between Armstrong's original patent application and the modern design is the stand-alone RF amplifier placed before the mixer. By World War II, some tube designs used frontend amplifiers comparable to today's FET frontends.

Mixer design structures today are highly similar to those of 100 years ago. However, owing to today's advances in materials, integration techniques, and materials disciplines, mixers based on diode and transistor structures have smaller sizes, higher sensitivities, higher bandwidths, and higher operating spectral ranges.

RF Mixers

Schottky-Diode Mixers

As early as the 1980s, American scientists developed GaAs Schottky-diode mixers with operating frequencies in the 1–10 THz range [7]. Subsequently, scientific research groups in the United States, Japan, the United Kingdom, and other countries launched research on Schottky diodes, including die designs, junction capacitance parallel effect, skin effect, plasma resonance, and other aspects [8].

The voltage and current of the Schottky diode have a nonlinear relationship, and the superconducting transformation can be completed without depending on the temperature conditions. It works at room temperature and has a low noise figure, wide operating bandwidth, and fast speed. It is an important diode mixer type at present. Commonly used structures in Schottky diode fundamental mixers include the cross-rod structure, ring bridge structure, and three-dimensional fin waveguide structure [9].

Schottky diode harmonic mixers can be divided into two circuit forms: monolithic integration and hybrid integration [10]. In the early stages, the working frequency band of high-order harmonic mixers was in the millimeter wave frequency

band. This was mainly because the high-frequency high-order harmonic mixer had higher requirements for the driving power of the local oscillator signal and was limited by the initial processing technology and research conditions.

Since then, domestic and foreign Schottky-diode mixers have been further developed, and hybrid integration technology has gradually matured and become practical. Diodes are attached to the substrate in hybrid-integrated harmonic mixers. Although some assembly errors are introduced, the mixers are widely used, owing to their simple process and low cost [11].

Common mixing structures are single-tube, single-balanced [12], and double-balanced [13]. In a simple Schottky diode-based mixing circuit, the harmonic components contained in the loaded local oscillator signal are mixed with the radiofrequency signal. The type of mixer can be used for both odd-order and even-order harmonic mixing. This structure has a simple design and small size, but poor isolation and noise figures; therefore, it cannot suppress clutter. To reduce frequency conversion losses, harmonic mixing circuits mostly use diode pairs in series or parallel.

Diode Types

Schottky Barrier Diodes

Schottky barrier diodes are widely used in the design of mixing devices. The key technology is the Schottky junction formed by the contact between the metal and N-type semiconductors, also known as the metal–semiconductor junction. Schottky junctions rely on carriers to transfer the current and have high electron mobility. With the advancement of modern semiconductor technology, Schottky contacts have been reduced to the order of microns. At present, diodes with cutoff frequencies of more than 10 THz have been developed, which can be used in terahertz circuit design.

As shown in Fig. 29.3, the diode is fabricated on an N-type semiconductor substrate with high electron mobility. Then, a high-purity, high-conductivity heavily doped N-type buffer layer is continuously grown on the substrate. This prevents substrate impurities from entering the epitaxial layer during the manufacturing process while ensuring a low series impedance. The buffer layer is generally doped with as many extra electrons as possible. For example, the doping concentration of GaAs is usually on the order of 10^{18} cm^{-3}.

The shallow-doped epitaxial layer continues to grow on top of the buffer layer, and its doping concentration is on the order of 10^{16}–10^{17} cm^{-3}. The metal anode contacts the epitaxial layer to form a Schottky junction; gold is often used to make air bridges (planar Schottky diodes) or whisker contacts (whisker Schottky diodes) [14].

Planar Schottky Diodes

Planar Schottky diodes first integrate the anode and cathode of the diode on the same metal substrate and then contact the metal anode with part of the barrier to form a semiconductor doped epitaxial layer, an N++ doped enhancement layer, and a metal cathode, in turn. With proper control, planar Schottky diodes can be manufactured entirely through photolithography, which greatly improves their repeatability. At

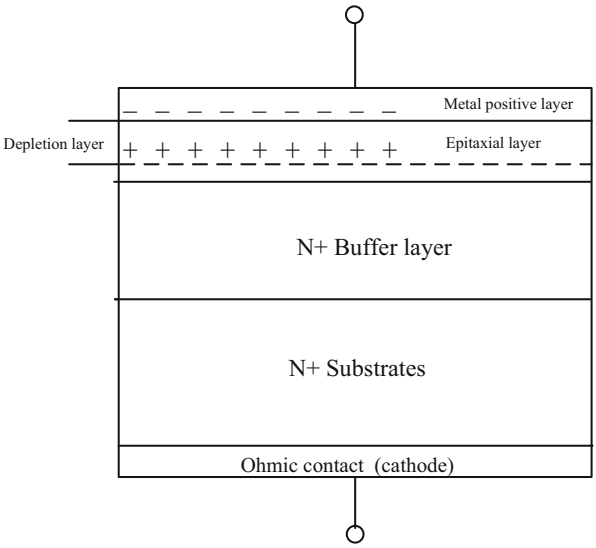


Fig. 29.3 Cross-sectional view of a Schottky barrier diode

present, two main types of planar Schottky diodes exist: quasi-vertical and surface channel [15].

The quasi-vertical Schottky diode improves the conductive structure using a tentacle contacting the Schottky diode. While retaining the same tube–honeycomb operation as in the Schottky-diode manufacturing process, crystal tentacles are applied using planar-circuit integration technology [16].

At present, the most mature planar technology is the surface channel planar diode. In recent years, surface channel diode technology has been rapidly developed in Lille, France, and the United Kingdom and has been applied to mixers [17, 18]. The physical structure of the surface channel diode is shown in Fig. 29.4. The parasitic capacitance introduced by this structure is greatly reduced, compared with the whisker structure Schottky diode, although compared with low-temperature superconducting mixing technology, its conversion loss is unsatisfactory [20].

Planar Schottky-diode development has gone through several core-technology innovations: the first was the development of flip-chip technology. In 2010, the GaAs substrate with high dielectric loss was removed, and the Schottky diode and its circuit were integrated. Using a quartz substrate with a low dielectric constant, the parasitic capacitance was further reduced [19].

In 2013, the Jet Propulsion Laboratory of the National Aeronautics and Space Administration (NASA) used a micro-integrated process to further generate a microstrip filter circuit, including a radiofrequency/local oscillator and intermediate frequency on a quartz substrate. This reduced the assembly error between the diode and surrounding circuits and removed the welding effect of the diode. In view of the original defect of upside-down diode technology—that is, the conductive glue overflowed the surface air channel and increased the parasitic capacitance when

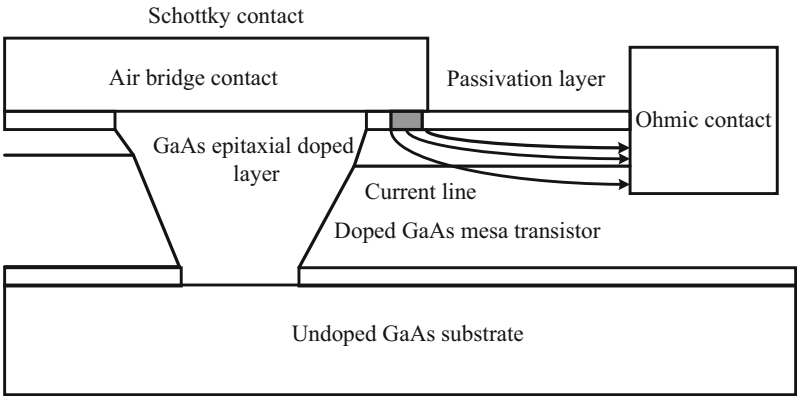


Fig. 29.4 Surface channel planar Schottky diode

pasted—the University of Virginia in the United States proposed conductive glue etching technology, which reduced the mixing loss of the mixer by 20–30% [20].

In contrast to the barrier material of traditional Schottky diodes, planar Schottky-diode barriers can be fabricated from several types of III-V materials, including gallium arsenide, indium gallium arsenide, or indium phosphide, or in combination, such as adding indium gallium arsenide to indium phosphide to make Schottky barriers.

If the semiconductor is to have high barrier performance, GaAs must be used to manufacture the Schottky barrier [21] because diode mixers made of other materials require a higher local oscillator drive power [22]. In addition, in the terahertz-frequency band, most semiconductor devices are made of GaAs, owing to its high electromigration rate and high breakdown voltage. For the above reasons, most planar Schottky diodes are currently made of GaAs and operate in frequency bands up to 600 GHz [23].

Diode Mixers

In 1971, Dominic A analyzed the nonlinear characteristics of Schottky barrier mixing diodes, produced a nonlinear model of them, and conducted a detailed analysis [24].

In 1973, Araki [25] designed a narrowband integrated balanced mixer. The mixer used two GaAs surface barrier diodes, and the circuit consisted of three couplers, a low-pass filter, and a transition joint from the waveguide to the microstrip. This circuit was representative of those for making microwave balanced mixers.

In 1977, McColl [26] analyzed the conversion loss of Schottky-diode mixers and explained in detail the effects of various factors, such as the diode junction capacitance, package-parasitic parameters, and impedance mismatch, on the conversion loss of mixers.

In 1977, Xie designed a fin-strip Schottky-diode W-band mixer. The frequency conversion loss of the mixer in the 85–95 GHz frequency band was less than 15 dB.

The fin-line coplanar-waveguide structure integrated mixer was studied accordingly and achieved a gain of 9 dB [27].

In 1978, Champlin et al. [28] extended factors, such as the skin effect, carrier inertia, and displacement current, to the traditional high-frequency model of non-epitaxial Schottky barrier diodes. The cutoff frequencies of the n-GaAs and n-Si extended models were calculated. Compared with the predicted results of the traditional model, both materials showed a wider operating frequency band.

In 1986, Adams et al. [29] proposed a new boundary condition for studying Schottky barrier diodes. Based on this modified boundary condition, the depletion of Schottky boundary electrons was predicted, and the current-dependent surface composite carrier-velocity distribution was studied. The electron-transport analysis was extended to higher bias voltages.

In 1993, Jelenski et al. [30] described GaAs Schottky barrier diodes with approximately ideal electrical and noise characteristics in terahertz mixing applications. The traditional formulas used to describe these characteristics were only valid within a limited forward-bias range, and the current was much smaller than the operating current under submillimeter mixing. The authors extended it to generalized analytical expressions of I-V and C-V properties of metal–semiconductor Schottky junctions over the entire bias range and proposed a new numerical diode model that occurred at the Schottky junction. This involved current-related recombination and drift phenomena, considering changes in the electron mobility and electron temperature in the unexhausted epitaxial layer.

In 1997, Lipsey et al. [31] observed that the harmonic circuit model of the diode equivalent-circuit model used to design high-frequency Schottky frequency doublers at that time could not accurately predict the characteristics of high-frequency circuits above 100 GHz. The physical principles of electronic transmission were analyzed.

In 2000, Hu et al. [31] developed a single-balanced mixer with a 93–96 GHz frequency and a conversion loss of less than 9 dB. In 2002, Dong et al. [32] developed a balanced mixer with a fin-coplanar structure of 92–96 GHz and a conversion loss of less than 15 dB; the port isolation was greater than 20 dB.

In References [51, 56], a fin-line balanced mixer with a 47-GHz structure was developed. The test results showed that the conversion loss was less than 3.5 dB, while the isolation between the local oscillator signal and the radiofrequency port could reach 20 dB. In 2008, Li Kan [33] developed a fourth-order harmonic mixer working in the Ka band, and its conversion loss was less than 11.8 dB at 34.36 GHz.

In 2003, Buffey et al. [34] developed a fin-line microwave mixer with a 350-GHz frequency, and its noise temperature reached 90 K. In 2005, Kyo Lee et al. [35] developed a new type of fin-line-coplanar line-structure balanced microwave mixer. The local oscillator power of the microwave mixer was as low as 6 dBm, the conversion loss reached 10 dB, and the signal isolation of the local oscillator port and the radiofrequency port was greater than 29 dB. In 2009, Thomas et al. [32] developed a sub-harmonic mirror rejection mixer. In its passband range of 320–340 GHz, its image rejection was as high as 7.2–28.1 dB.

In 2009, Siles et al. [36] believed that the traditional boundary conditions of the previous drift diffusion model at the Schottky contact were limited and could not

correctly predict the working characteristics of the Schottky-diode mixer. This resulted in large differences between the simulation and the actual measurements. Therefore, a Monte Carlo simulator was used to study the working characteristics of Schottky diodes in flat band mode, which improved the Schottky boundary conditions, enhanced the validity of the drift diffusion model, and accurately predicted the mixer's state in the flat band state. The diode model considered factors, such as barrier lowering, tunneling, carrier-velocity saturation, and mobility field-related effects.

In 2011, Chao et al. [37] developed a Schottky barrier varactor diode that could be applied to microwave-frequency doubling circuits using standard microelectronic technology. The planar-structure manufacturing process was used to overcome the difficulty of traditional manufacturing process integration. The doping concentration in the N-type layer was exponential, which made the varactor ratio higher than that of the traditional uniform-doping structure. This improved the operating frequency and output power of the frequency doubling circuit.

In 2012, Saber [38] conducted a detailed study of the thermal properties of terahertz Schottky diodes based on self-heating effects. The thermal distribution of diodes at different physical sizes was analyzed, including the maximum junction temperature, thermal resistance, and thermal time constants.

In 2013, Yan et al. [39] proposed a Schottky barrier diode based on physical parameters, such as anode area, ohmic contact area, doping distribution of epitaxial layers and substrate layers, thickness, barrier height, contact resistance, and device temperature. The physical base model of the diode considered other factors, such as the reduction of the barrier height, owing to the mirror force of free electrons close to the barrier, the nonlinear resistance of the epitaxial layer, and thermal electron noise. The physical base model of the diode was applied to a reverse-parallel Schottky diode. The simulation results were basically consistent with the measured results of the I-V and C-V characteristics.

In 2014, Pérez-Moreno et al. [40] proposed a physical-based Schottky-diode numerical electrothermal model that characterized the effect of self-heating on the performance of terahertz inverter devices and circuits. The model consisted of a drift diffusion equation enhanced by a heat-conduction equation, which considered the interaction between electrons, holes, and lattice subsystems in semiconductor devices. Introducing lumped thermal resistance, a thermal resistance matrix of diodes was established to characterize the thermal interaction between anodes. The established electrothermal model was integrated into a harmonic balance circuit simulator, and the Schottky-diode circuit was analyzed, designed, and optimized. The simulation results of the 200-GHz frequency multiplier based on the diode physical base model were consistent with the measured results at high input power.

In 2020 and 2021, Orfao et al. [41] used two-dimensional (2D) integrated Monte Carlo simulations to study the surface charge and edge capacitance effects of GaAs and GaN Schottky barrier diodes. The edge effect parameters were defined, and the effects of process parameters, such as the insulating material, insulation thickness, and epitaxial layer thickness, on the C-V characteristics of Schottky diodes were analyzed. Parameters, such as the insulating material used

for passivation, its width, and the lateral extension of the epitaxial layer, significantly affected the edge capacitance, thereby increasing the total capacitance value above the ideal value. The diode edge effect of high-dielectric-constant insulating materials was more obvious and strengthened the electric field extending laterally from the epitaxial layer. Compared with the models of foreign countries, the research level and depth of domestic terahertz Schottky-diode models have great room for improvement [42].

Owing to the complex physical mechanism of Schottky diodes in the terahertz frequency band, the frequency conversion circuit places high requirements on model accuracy. Therefore, it is necessary to deeply explore the physical mechanism of Schottky diodes in the terahertz frequency band, establish an accurate Schottky-diode model, guide the manufacture and design of Schottky diodes, and lay a solid foundation for creating high-performance terahertz frequency multipliers and mixers.

CMOS Mixers

CMOS mixers can create high-performance integrated modular circuits with low cost and high integration. They have become a research hotspot in RF integrated circuit design. RF circuit structures and design technology based on the CMOS process have been proposed. Gradually, mature CMOS mixer products and designs have been introduced and have achieved great results.

In 1984, Tang designed a CMOS switching circuit mixer [43]. This mixer was completely different from the then-current CMOS mixer. It used four analog–digital bidirectional switches to do the mixing, instead of using an integrated mixer chip. The analog–digital switch of this circuit had a limited frequency, could not be used at high frequencies, and had a large conversion loss.

In 1996, foreign research on high-linearity mixers began. Toole, El-Masry, and Manku designed a high-linearity mixer. It was a subsampled mixer, which is different from the mixers commonly used today. Although its other performance metrics were poor, the linearity was very high. The power of the third-order intermodulation component signal was as high as 35 dBm, and the power consumption was as low as 2 mW [44].

Since 1998, the performance indicators of CMOS mixer chips have been increasingly optimized, and mixer designs have begun to diversify. Giustolisi et al. proposed a mixer with a low power supply voltage in 2018. Although the common active double-balance structure was used, the power-supply voltage was reduced from 5 V and 3 V to 1.5 V, and the power consumption was relatively low, only 3.4 mW [45].

In 1998, Wang proposed a low supply-voltage mixer, which only required 1 V to work normally. The lowest power consumption was only 0.14 mW. Li et al. analyzed the noise in the mixer in 2019 [46–48].

Since 2003, research on CMOS mixers has increased, design ideas have matured, and integrated chip mixer designs have begun [49]. Shi and Chi designed a down-conversion mixer operating at a 2.4-GHz frequency with a 0.18-mm CMOS process in 2003. The mixer was active and used the most commonly used double-balanced

structure [50]. Although its performance seems low now, it is still very valuable for learning [45].

In 2003, El-Shewekh and El-Saba proposed a broadband mixer with an RF input range from 30 MHz to 1 GHz, a port isolation power of up to 29.5 dBm at a gain of up to 29.1 dB, with low noise and power consumption [51].

In 2004, Cui used a 0.35-mm CMOS process to design up-conversion and down-conversion mixers working at 2.4 GHz [49]. Both mixers used an active double-balanced structure and were improved on this basis. In addition, this study revealed one of the development trends of future mixers: the power-supply voltage consistently decreased [52, 53].

In 2005, Dong proposed a new design structure for ultra-wideband mixers. The RF input range was approximately 3.1–5.15 GHz, the output was close to 1 GHz, and its bandwidth was very wide [52]. The development trend of future mixers is increasingly wide bandwidth [54, 55].

In October 2005, Wei et al. analyzed the linearity of a mixer based on a low power supply voltage, and the port-output linearity reached 7.4 dBm [56]. It revealed that future mixers will have increasingly high linearity [57, 58].

In 2006, Luo et al. analyzed the noise of active mixers [59], analyzed the source of the noise in active mixers, and studied the influence of factors, such as the transconductance stage, switch tube, and local oscillator signal size, on the noise. Low-noise mixers have also received attention and research [60, 61].

As mixers have been developed, special mixer structures have appeared. In 2010, Yuan proposed a current-mode mixer. Unlike voltage-mode mixers, current-mode mixers mix current signals. Compared with traditional voltage-mode mixers, current-mode mixers are more suitable for low-voltage, high-linearity designs [62].

Beginning in 2012, low-power mixers [63] began to increase, and since then, the power-supply voltage of low-voltage mixers has been lower than 2 V [63]. In 2012, Shen and Zhang proposed an ultra-low-power mixer for aerospace applications. The power consumption was only 0.76 mW and the power-supply voltage was only 0.6 V. This power consumption and power-supply voltage are also very low today [64].

Since 2013, owing to advancements in the process, the operating frequency of CMOS mixers has increased. Unlike the previous several GHz, the operating frequency of mixers at that time could reach tens or even hundreds of GHz. In 2013, Wang used a Schottky barrier diode to design a mixer that worked at a frequency of up to 340 GHz [65]. In 2000, Terrovitis and Meyer analyzed the linearity of a mixer and provided a theoretical basis for the subsequent improvement of high-linearity mixer designs [65].

Mixer Circuit Structures

Single-Ended Structures

Unbalanced mixer circuits are more commonly used; the circuit is usually composed of a single transistor, FET, or double-gate FET. Its principle circuit diagram is shown in Fig. 29.5. The information-carrying signal and the local

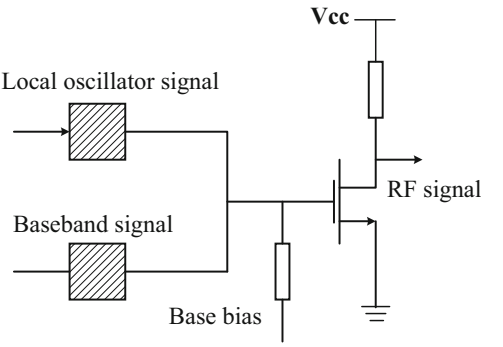


Fig. 29.5 Unbalanced mixer [66]

oscillator signal are simultaneously received by the base of the transistor, and the RF signal is output at the collector. The local oscillator signal and the transistor bias voltage are used together as the bias signal. At this time, the transconductance g_m of the transistor becomes a time-varying transconductance.

If the baseband input signal is V_{BB}, the output current of the mixer can be expressed as

$$i_{RF}(t) = v_{BB}(t) \cdot g_m(t) \approx \frac{1}{2} g_m V_{BB} \cos(\omega_{LO} - \omega_{BB})t.$$

The unbalanced mixer has a simple structure and therefore has good noise performance. However, because the local oscillator signal is unbalanced, the mixer cannot suppress the signal crosstalk from the local oscillator port to the radiofrequency port; thus, the isolation between the ports is low [66].

Single-Balanced Structures

The most commonly used structures in the antiparallel diode pair (APDP) circuit structure in a single-balanced mixer include the microstrip-branch type [67], duplex-tube-pair parallel type, and duplex-tube-pair series type [68]. The structures are shown in Fig. 29.6.

Figure 29.7 shows the structure of a single-balanced active mixer. It consists of a baseband signal-driven transconductor and a local oscillator signal-driven differential pair.

Compared with the unbalanced mixer, the single-balanced mixer reduces the leakage of the local oscillator signal to the radiofrequency signal, owing to the balanced input of the local oscillator signal, and improves the isolation between the mixer ports. Moreover, owing to the characteristics of the switch tube, it can achieve good switching-performance linearity at a lower local oscillator driving voltage and reduce the system power consumption.

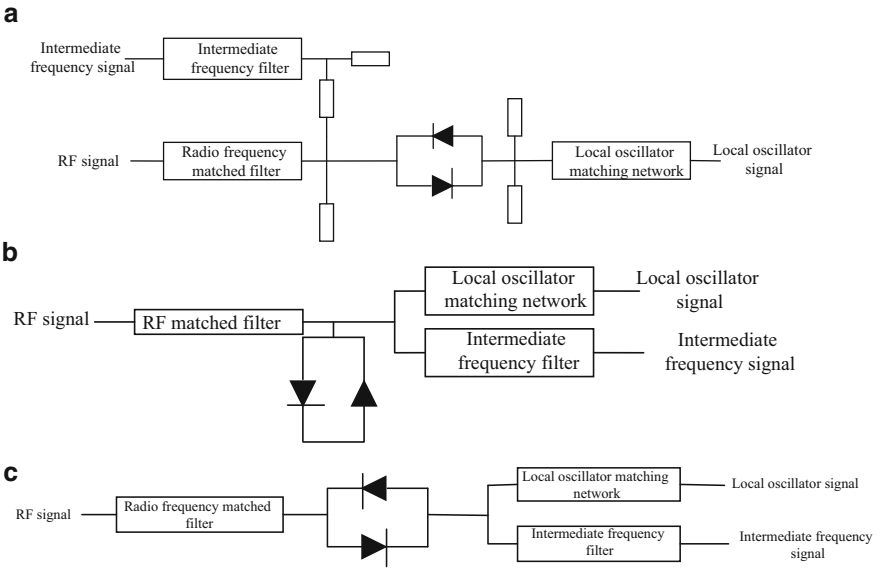


Fig. 29.6 Schematic diagram of typical mixer structures [20]. (**a**) Microstrip-branch structure, (**b**) duplex-tube-pair parallel structure, (**c**) duplex-tube-pair series structure

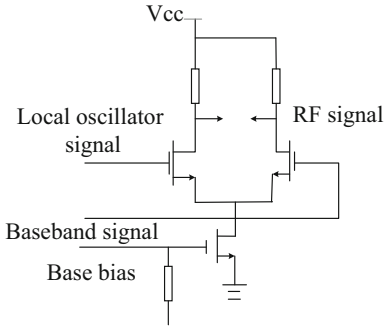


Fig. 29.7 Structure of a single-balanced active mixer

Double-Balanced Structures

Combining two single-balanced mixers so that the baseband signal is also input by differential means constitutes a double-balanced mixer. The Gilbert double-balanced mixer is the most common double-balanced mixer in integrated circuits. As shown in Fig. 29.8, it mainly consists of two differential pairs of six transistors: two transistors form a differential structure, which converts the input baseband signal into current; four transistors are each turned on under the driving of a large local oscillator signal, and then, the current converted by the baseband signal is modulated.

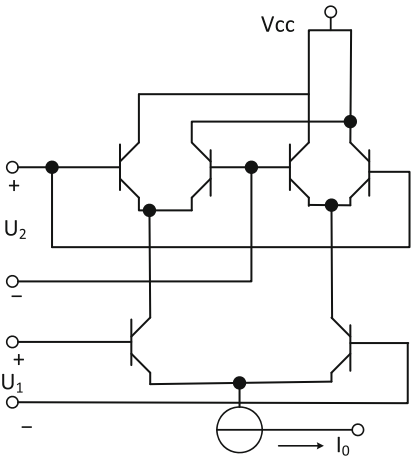


Fig. 29.8 Gilbert double-balanced mixer [69]

The Gilbert double-balanced mixer mainly has the following advantages:

1. Small footprint and low power consumption.
2. The output can effectively suppress DC, local oscillator, and other odd harmonic components.
3. It can use relatively small local oscillator power in the form of differential switching.
4. It adopts a double-balanced structure, which has higher isolation than other structures while ensuring certain linearity, and is very suitable for direct-conversion transmitter designs.
5. It has a good yield in the production of integrated circuits.
6. Therefore, it is widely used in radiofrequency integrated circuits.

Unfortunately, the linear response range of the traditional Gilbert mixer is relatively low, and noise analysis is difficult. Specifically, it adopts a three-layer transistor structure, which makes it difficult to achieve a low voltage and increases the system power consumption.

Research on double-balanced mixers has been mature since the 1980s, with many reports in the literature. Hybrid-integrated millimeter-wave mixers in the form of fin lines, coplanar lines, and suspended microstrips have developed rapidly. In 1982, Howe et al. developed a novel broadband double-balanced mixer with a frequency range of 21–40 GHz. The design balanced the microstrip-line coupling signal and local oscillator signal by balancing the slot line and the branch line to the beam-lead Schottky diode encapsulated in a star structure [70].

Triple-Balanced Structures

In the 2010s, triple-balanced mixer structures could be used to make wider-band mixers, which were favored because of their advantages, such as high dynamic range and intermodulation rejection [71]. In 2010, Ton et al. developed an X-band monolithic dual-balanced mixer with a third-order intermodulation input of 20 dBm. When the local oscillator power was 20 dBm, the output power of its 1-dB compression point was 3.6 dBm. In the range of 7–10 GHz, the conversion loss was less than 10 dB, and the intermediate frequency was 5 GHz. This high-input third-order intermodulation mixer was a key technology in high dynamic range receivers, and the successful development of this technology greatly promoted the development of integrated monolithic receivers [72].

The development of GaAs processes and monolithic technology has pushed research on submillimeter-wave and millimeter-wave integrated mixing circuits to a new height. In 2015, John et al. invented a new planar double-double-balanced monolithic-microwave integrated circuit (MMIC) mixer structure. The structure used a mixture of slot lines, coplanar waveguides, and coplanar lines and was mounted on a substrate with eight Schottky diodes. The mixer implemented with this structure had a radiofrequency of 6–20 GHz, a local oscillator frequency of 8–18 GHz, an intermediate frequency of 2–7 GHz, and a conversion loss of 6.2–9.8 dB. The successful development of this mixer laid the foundation for the development of planar-structure double-balanced mixers [73].

Mixer Linearity

The linearity of the mixer directly determines the dynamic range of the receiver; therefore, improving the linearity has become a key topic in the mixer research field. The main reason for the third-order distortion of active mixers is the nonlinearity in the transconductance stage. Consequently, research work on mixer linearity mainly focuses on the problems related to the transconductance stage of the mixer.

In 2002, Salem et al. proposed a radiofrequency input technology without a tail current, which greatly improved the linearity of the mixer; however, the structure has a low common mode rejection ratio [74]. Another common method to improve linearity is the source-level inductance negative feedback method.

In mixer-linearization research, the main linearization techniques are as follows:

1. Class AB transconductance stage. In 2004, Wei et al. proposed a mixer composed of class AB transconductance stages. This method improved the linearity of the mixer; however, using the transconductance stage increased the power consumption of the mixer. Moreover, spurious components were generated in the output stage [75].
2. Piecewise linear technology. Shen et al. [76] adopted piecewise linear technology, which uses multiple differential pairs with independent base bias voltages in parallel. Because each differential unit is linear in the input range, many such

differential pairs are reasonably moved along the input, and the output appears as the superposition of multiple linear intervals, generally obtaining a large linearity. However, as the number of differential pairs increases, the power consumption and input capacitance will increase accordingly.

3. Feedback technology. Aparin et al. [77] used feedback technology to improve the linearity. Observing the DC transfer characteristics of field effect transistors, the third-order derivation changes from positive to negative in the inverse region, which is the main source of third-order nonlinearity. Feedback technology can improve linearity by canceling this nonlinear effect. However, the transistor must be biased at just the right static operating point, which limits the transconductance of the input stage, resulting in a decreased gain and increased noise figure.
4. Derivative superposition (DS) and improved DS technology. Webster et al. [78] first proposed the concept of DS technology. It is realized by connecting in parallel two transistors working in the weak and strong inversion regions, respectively. By carefully selecting the width and bias voltage of the transistors, the positive third-order derived peaks working in the weak inversion region and the negative third-order derived peaks working in the strong inversion region cancel each other out. In this manner, the third-order derived term will be close to zero, resulting in a greatly improved linearity.

   Reference [79] adopts this method to increase the low noise compensation amplifier from 5.5 dBm before compensation to 15.6 dBm, while the frequency cutoff point of the third-order harmonic signal of the mixer is increased from 2.5 dBm before compensation to 9 dBm. Because the auxiliary metal oxide semiconductor field effect transistor (MOSFET) works in the weak reflection region, this method does not increase the power consumption of the mixer.

   Traditional DS technology only considers how to eliminate the third-order transconductance; it does not consider the influence of second-order transconductance on the circuit. When the third-order transconductance of the circuit is close to zero, the second-order transconductance becomes the main factor that produces the third-order nonlinearity [79].

   Reference [20] proposes an improved derived term superposition technique that considers the effects of the second-order nonlinearity. This method uses two inductors to tune the second-order and third-order nonlinearities that have a major impact on the third-order intermodulation term. The amplitude and phase of the order nonlinearity make the amplitudes equal and phases opposite. Thus, the M_3 term is zero at the output, and the port isolation of the low-noise amplifier (LNA) is increased from 10.5 dBm before compensation to 19.2 dBm [80].
5. Current injection method. A current source is added to the common source node of the local oscillator stage to increase the DC current of the transconductance stage without affecting the current of the local oscillator stage. Thus, the linearity and gain are increased without affecting the noise, producing good linearity and noise characteristics [81].

Optical Mixers

Optical communication technology modulates the information to be transmitted onto a beam of light and transmits it in a channel. Optical communication can be divided into two ways, according to the modulation method at the transmitting end: incoherent optical and coherent optical.

Incoherent optical communication uses high-frequency light intensity modulation to load information onto the light beam and receives the transmitted optical information through direct detection. Its transmission bandwidth is limited and the detection sensitivity is low. Therefore, in addition to modulating the light intensity, coherent optical communication modulates the optical amplitude, frequency, and phase, with various modulation and demodulation methods at the receiving end. The heterodyne-receiving sensitivity is about 20 dB higher than that of the direct detection method and has good filtering performance.

The receiving equipment is easy to integrate and is suitable for long-distance and large-capacity communication. The system is mainly composed of three parts: a modulation transmitter, transmission channel, and digital coherent receiver. The coherent receiver is composed of an optical mixer and a balanced detector. The received optical signal and local oscillator signal are input to a 90° optical mixer for mixing, and the balanced detector completes the photoelectric conversion.

Because the optical mixer is the main component of the digital coherent receiver, its performance directly determines the coherent receiver signal-to-noise ratio and reception sensitivity, which affect the communication quality of the entire coherent optical communication. Therefore, optical mixer research mainly improves the digital coherent receiver. The key to sensitivity is also a key research area in the field of coherent optical communication.

Working Principle

The optical mixer interferically mixes the received signal light and the local oscillator light and outputs the mixed optical signal. The balanced photodetector detects the mixed optical signal and outputs the intermediate frequency current signal, and the demodulation circuit demodulates the signal to extract the original signal. In addition to its wide application in the field of coherent optical communication, optical mixing has many applications in photon imaging [81], lidar [82], biomedical sensing [83], and other fields.

As shown in Fig. 29.9, the main function of the optical mixer is to accurately guide the input signal light, interfere with the local oscillator light, split the light, and change the phase. Finally, the intermediate frequency signal is output and demodulated through balanced detection.

Current optical mixers can be divided into two categories according to their output forms: two output 180° optical mixers and four output 90° optical mixers. Most coherent communication receiving systems use a 90° optical mixer, which is a six-port device with two ports for input and four ports for output [84].

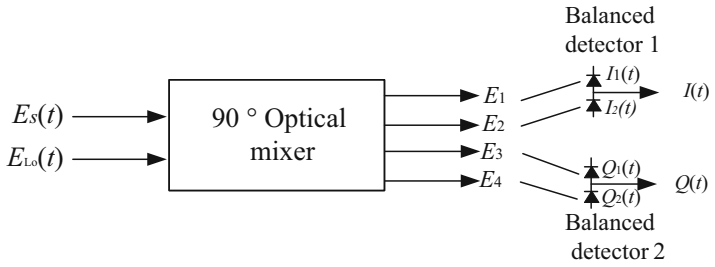


Fig. 29.9 Schematic diagram of the working principle of the 90° optical mixer

The signal light and local oscillator light can be expressed as

$$E_s(t) = \sqrt{P_s}\exp(j\omega_s t + j\varphi_s(t)) \tag{29.1}$$

$$E_{LO}(t) = \sqrt{P_{LO}}\exp(j\omega_{LO} t + j\varphi_{LO}(t)). \tag{29.2}$$

At the receiving end of coherent optical communication, the signal light is first mixed with the local oscillator light, where P_s and P_{LO} are the optical powers of the signal light and local oscillator light, respectively, w_s and ω_{LO} are their angular frequencies, respectively, and $j_s(t)$ and $j_{LO}(t)$ are their initial phases, respectively. Their phase difference $\Delta\phi(t) = \varphi_s(t) - \varphi_{LO}(t)$, that is, the information to be demodulated. The signal light and the local oscillator light pass through an ideal 90° optical mixer to obtain the output optical signal:

$$E_1 = \frac{1}{2}(E_S + E_{LO}) \tag{29.3}$$

$$E_2 = \frac{1}{2}(E_S - E_{LO}) \tag{29.4}$$

$$E_3 = \frac{1}{2}(E_S + jE_{LO}) \tag{29.5}$$

$$E_4 = \frac{1}{2}(E_S - jE_{LO}). \tag{29.6}$$

After being detected by the balanced detector, the corresponding four output electrical signals can be obtained:

$$I_1(t) = \frac{R}{4}\left(P_s + P_{LO} + 2\sqrt{P_s P_{LO}}\cos(\Delta\omega t + \Delta\phi(t))\right) \tag{29.7}$$

$$I_2(t) = \frac{R}{4}\left(P_s + P_{LO} - 2\sqrt{P_s P_{LO}}\cos(\Delta\omega t + \Delta\phi(t))\right) \tag{29.8}$$

$$Q_1(t) = \frac{R}{4}\left(P_s + P_{LO} + 2\sqrt{P_s P_{LO}}\sin(\Delta\omega t + \Delta\phi(t))\right) \tag{29.9}$$

$$Q_2(t) = \frac{R}{4}\left(P_s + P_{LO} - 2\sqrt{P_s P_{LO}}\sin(\Delta\omega t + \Delta\phi(t))\right), \tag{29.10}$$

where R is the response of the detector and $\Delta\omega = \omega_s - \omega_{LO}$ represents the angular frequency difference between the signal light and the local oscillator light. Using zero difference detection, $\Delta\omega = 0$. The current values of the obtained I channel and Q channel are

$$I(t) = I_1(t) - I_2(t) = R\sqrt{P_s P_{LO}}\cos(\Delta\phi(t))$$

$$Q(t) = Q_1(t) - Q_2(t) = R\sqrt{P_s P_{LO}}\sin(\Delta\phi(t)).$$

After subsequent digital signal processing, the original signal and the information to be transmitted can be extracted [85].

Research Progress and Current Status

Optical mixers can be mainly divided into optical fiber, free space, and planar waveguide types. In addition, there are many structures in different types of mixers. Optical mixer structures have several important types, such as the 3-dB coupler-type, multimode interference (MMI) coupler-type, crystal refraction type, and waveguide technology [86]. According to the coupling method, optical mixers can also be divided into optical fiber and spatial optical coupling types [87]. According to the integrated substrate, they can be divided into silicon dioxide substrate types and polymer waveguide base types [88]. Next, we introduce optical mixers of different coupler and base types and discuss their applications in the field of coherent optical communication.

3-dB Coupler and Phase Shifters

A widely used optical mixer structure combines a 3-dB coupler and a phase shifter. This structure can output the mixed optical signal in a specific polarization state and can be integrated and prepared on lithium niobate material or silicon dioxide material. It has a small output phase difference and stable performance.

In 1983, Yuen first proposed using a polarization splitting prism to mix two beams of coherent light [89] and extracted information through coherence. This was the prototype of the early concept of optical mixers. Subsequently, in 1985, Abbas added filters and noise amplifiers to the Yuen simple-mixer structure [90]. It could simultaneously filter the local oscillator noise and amplify weak signals.

In 1987, Gimlett [91] designed a 90° optical mixer based on a 3-dB fiber-optic coupler with polarization beam splitters (PBS), as shown in Fig. 29.10, to achieve

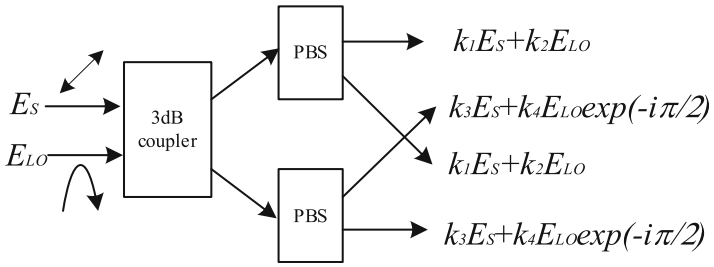


Fig. 29.10 Schematic diagram of an optical fiber 3-dB coupler type mixer [95]

frequency shift keying (FSK) at a rate of 2 Gbit/s. When the bit error rate was 10^{-9}, the sensitivity of the coherent receiver reached −36.7–39.2 dBm.

In 2010, Nara [92] used a directional coupler instead of a 3-dB coupler to couple light beams and combined the structure with a phase shifter to make a 90° optical mixer using silicon dioxide. The insertion loss of the device in the C band was less than 10 dB, the polarization extinction ratio exceeded 20 dB, and the phase deviation was less than 1°.

In 2021, Chen et al. [93] designed a planar waveguide optical mixer using a combination of a 3-dB coupler and a 90° phase shifter. The mixer used silicon dioxide with a high refractive index difference as the base material to improve the phase shifter. The waveguide was increased to 2.2 to 3.6 times that of the length difference scheme. In the experiment, the phase error of the optical mixer was less than 2°, the loss at the same phase end was less than 1.18 dB, and the insertion loss of the entire optical mixer was less than 1.22 dB.

The optical mixer based on the 3-dB coupler and phase shifter has a relatively simple structure, low manufacturing process requirements, and relatively mature manufacturing technology. Therefore, it has the advantages of a small phase deviation, strong adaptability to the environment in practical applications, and stable working performance. It is widely used in optical-fiber communication and space optical communication; however, the device size is large, and the characteristics of the silicon-based platform and 3-dB coupler structure make it inconvenient to integrate.

Mach–Zehnder Interference Optical Mixers

Mach–Zehnder Interference Optical Mixer Structure

An optical mixer with a Mach–Zehnder interference structure generally consists of three parts. The schematic diagram of the structure is shown in Fig. 29.11. The first part is a coupling region with a length of L_1, which produces a coherent beam of light with a certain amplitude ratio and phase difference. The second part is a non-coupling region, which introduces a length difference of ΔL in one arm to generate a phase difference related to the wavelength; it then adjusts the phase

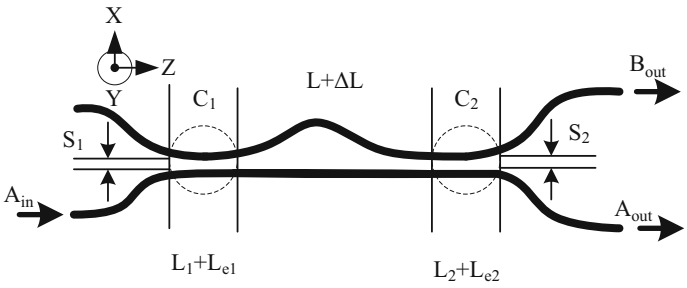


Fig. 29.11 Principle of a Mach–Zehnder interference mixer [94]

difference of the two beams of light, that is, performing phase compensation. The third part optimizes the coupling ratio region to control the coupling ratio of the output beam. By simultaneously optimizing the lengths of these three parts, wavelength independence for different coupling ratios can be achieved [94].

The expression for coupling efficiency is as follows:

$$\eta = \frac{|B_{\text{out}}|}{|A_{\text{out}}| + |B_{\text{out}}|} = a^2 + b^2 + 2ab\cos(\beta \Delta L), \tag{29.11}$$

where h is the coupling efficiency, A_{out} and B_{out} are the two output power signals, respectively, β is the refractive index coefficient of the mixer, and ΔL is the length of the mixer. Equation (29.11) is the mixing-efficiency formula of the Mach-Zehnder interferometric mixer, wherein parameters a and b are related to the length of the interference cavity. The specific expressions are shown in Eqs. (29.12) and (29.13), respectively.

$$a = \cos\left[\frac{\pi}{2L_{\text{c1}}}(L_1 + L_{\text{e1}})\right] \sin\left[\frac{\pi}{2L_2}(L_2 + L_{\text{e2}})\right] \tag{29.12}$$

$$b = \sin\left[\frac{\pi}{2L_{\text{c1}}}(L_1 + L_{\text{e1}})\right] \cos\left[\frac{\pi}{2L_2}(L_2 + L_{\text{e2}})\right] \tag{29.13}$$

Mach–Zehnder Interference Optical Mixing

In 2003, Chen [95] used the waveguide of a fluorine-doped polymer material (polyimide) to theoretically achieve a coupling ratio that fluctuated by 0.8% in the 160-nm wavelength range in the Mach–Zehnder interference optical mixer. It also achieved polarization insensitivity.

In 2005, He et al. [96] developed an optical mixer structure, based on the Mach–Zehnder structure and using a genetic algorithm. They theoretically achieved a coupling ratio in a silicon dioxide waveguide in the 200-nm wavelength range with a $5 \pm 1.5\%$ coupling ratio.

In 2013 and 2014, Alam [97] improved the Mach–Zehnder structure, using a hybrid plasma-waveguide structure for phase compensation, and implemented a compact large-bandwidth Mach–Zehnder interferometric optical mixer in a silicon-on-insulator (SOI) waveguide.

In 2010, Zhou et al. [98] designed a 2×4 90° optical mixer for free space optical communication and applied a zero difference interference-detection receiver, the basic structure of which is shown in Fig. 29.12. The asymmetric Mach–Zehnder interference mixer split the signal light into two channels, interfered with the local oscillator light, and recombined the two signals. The balanced detector received the mixed light and converted it into an intermediate frequency current output. The contrast ratio of the interferometer ground modulation or signal-to-noise ratio depends on the wavefront interference process, rather than the wavefront quality, so no wavefront compensation technology or frequency locking is required. This scheme effectively eliminates the influence of wavefront distortion on the coherent detection process.

Optical mixer schemes, such as waveguide, 3-dB coupler-type, and prism-type optical mixers, only have a down-conversion function. In practical applications, the optical mixer must have simultaneous up-down conversions. Therefore, an optical

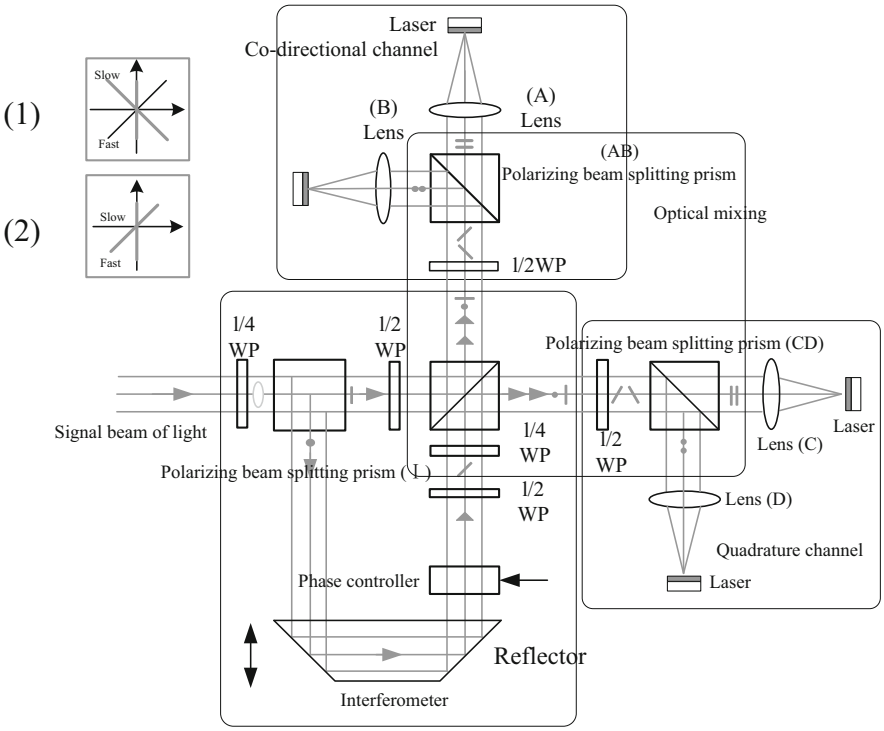


Fig. 29.12 High-precision 90° spatial optical mixer [98]

mixer was designed, using the down-conversion principle realized by Mach–Zehnder interference modulation [99].

In 2021, Zhang et al. [100] proposed an optical mixer design scheme using a dual Mach–Zehnder interference modulator. In this scheme, the signal light and the local oscillator light, respectively, realize the local oscillator light and the signal light in the Mach–Zehnder interferometer. Single-sideband modulation modulates the phase difference between the upper and lower arms of the Mach–Zehnder interference, so that the optical carriers cancel each other, and outputs the intermediate frequency signal of the optical signal or the up-conversion signal. This scheme realizes up-down conversion and further streamlines the transceiver module of a coherent optical communication system.

Polarized Optical Mixers

Polarized Optical Mixer Structure

Garreis et al. [101] designed a 90° space optical mixer with a block structure in the late 2010s. This scheme has been used many times in the laser communication end points developed by NASA and the European Space Agency. The principle is shown in Fig. 29.13.

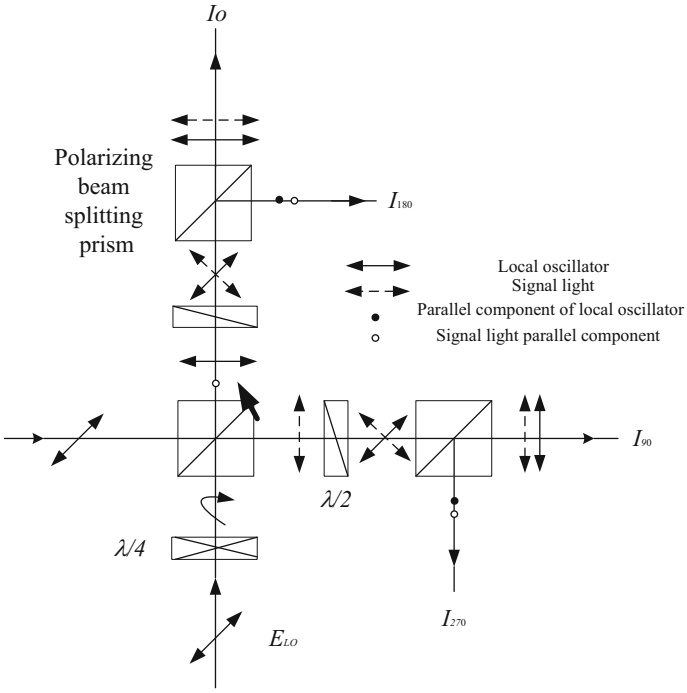


Fig. 29.13 Schematic diagram of a polarized light mixer [101]

The stability of the laser performance, channel conditions, encoding method, randomness of atmospheric transmission, platform vibration, etc., will all affect the polarization state of the signal light. The interior of a spatial light mixer mostly contains polarization-sensitive optical elements. The mismatch and change of the polarization states of the signal light and the local light seriously affect the intensity and phase relationships of the output mixed light, which, in turn, limit the mixing efficiency of the spatial light mixer, and even affect the detection efficiency of the entire coherent optical communication system for signals [102]. Compared with linearly polarized light, circularly polarized light and elliptically polarized light have stronger adaptability to the atmosphere, which helps improve the certainty of the polarization state of the communication light.

The polarization characteristics of the light beam introduce compatibility problems to the working mode of the integrated device and the fiber-input mode. This is because the light mode is affected by other factors and changes randomly. Silicon-based integrated devices are polarization dependent, so the device performance is not guaranteed. Polarization control is necessary to take advantage of the advantages of polarization sensitivity or avoid its disadvantages [103].

The basic principle of polarization diversity [104] is as follows: First, the input light is divided into two mutually orthogonal polarization states, transverse-electric (TE) mode and transverse-magnetic (TM) mode, after passing through a polarizing beam splitter (PBS). The polarization rotator converts the light of one of the polarization states into an orthogonal polarization state and then inputs it into two identical silicon-based integrated devices; then, one of the output lights is converted back to the initial polarization state. Finally, the two orthogonal polarization states are combined into a light beam through a polarization beam combiner. Therefore, a large number of polarization splitting prisms and optical mixer structures based on them have appeared for the polarization characteristics of light beams.

Research Progress of Polarization Optical Mixers

In 1968, Teich [105] of the Massachusetts Institute of Technology (MIT) designed an optical mixer based on a beam splitter, which obtained an intermediate frequency signal by the coherent superposition of a signal light and local oscillator light.

In 1969, Kleinman of Bell Labs [106] used a polarizing beam splitter to transmit and reflect light beams of two specific wavelengths and then realized an optical mixing function. The working principle is shown in Fig. 29.14. The system uses a polarization splitting prism, a visible light laser, and a photomultiplier tube to detect the optical signal generated by the sum or difference frequency. The signal light and the local light pass through the polarization splitting prism to produce interference mixing, and the balanced detector receives and converts it into an electrical signal.

In 1989, Langenhrost et al. [107] designed a compact block optical mixer, which used a gradual-change refractive index lens to shape the input and output light beams, so that the input beam was quasi-straight and the output beam was focused. The polarization splitting prism coupled the signal light, and the calcite crystal separated the beams according to the polarization state. The insertion loss of the mixer was only 2 dB, and the average coupling efficiency reached 63%.

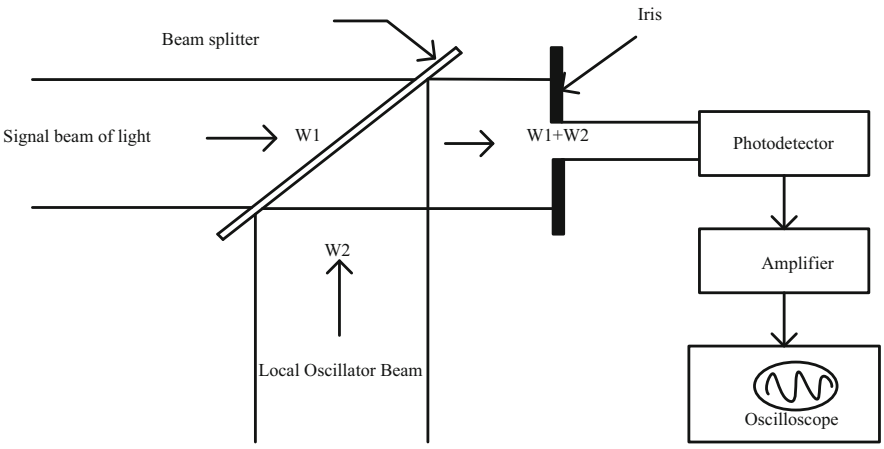


Fig. 29.14 Spatial light mixer using a beam splitter [106]

In 2011, Garreis [108] proposed a six-port 90° spatial optical mixer composed of a polarized spectroscopic prism and a non-polarized spectroscopic prism. The mixer had a simple structure, high stability, and easy processing. The structure has been used by TESAT many times.

In 2016, Zhang et al. [109] designed a polarized 90° spatial light mixer. The optical path of the mixer is shown in Fig. 29.15. After the signal light is incident, it interferes and mixes with the local oscillator light. The signal light after mixing passes through different waveplate units to generate four optical signals with a 90° phase difference. Considering the special needs of optical mixers in laser communication systems, a 90° mixer with spatial light input and local oscillator light input using a polarization maintaining fiber was designed and completed.

In 2017, Zhang et al. [110] designed a 90° spatial optical mixer with adjustable branch energy. By controlling the incident ray polarization state, it controlled the reasonable distribution of optical energy between the communication branch and the phase locked branch to improve the system performance. In addition, a 90° spatial mixer using a single crystal was also designed (Fig. 29.16).

In 2013, Kawashima [111] designed a polarization diversity 90° spatial light mixer, which used a polarization beam splitter and waveplate to mix the local oscillator light with the signal photosynthetic beam. It then split the beam to adjust the phase, so that the four output signals had a 90° phase difference, to achieve coherent mixing. In the actual prototype, the four output lights all use optical fiber as the output interface, as shown in Fig. 29.18.

In addition to the waveplate structure, there are also applications with polarizing beam splitters and Mach–Zehnder interference modulators. The light loss of these optical structures is small, the contrast of the interference modulation is high, and the interference technology is based on the phase difference, rather than the beam. The wavefront quality of light thus greatly reduces the impact of wavefront distortions on the mixing efficiency.

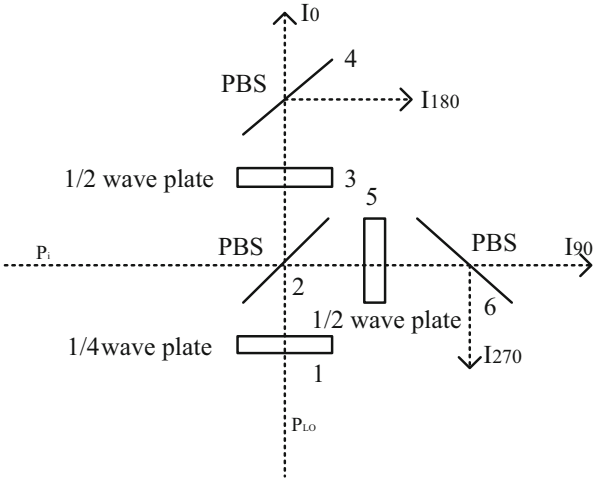


Fig. 29.15 Optical path diagram of a polarization type optical mixer [109]

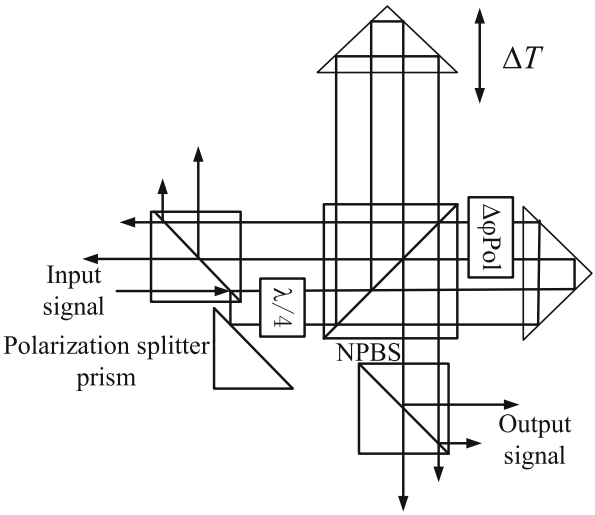


Fig. 29.16 MIT polarization diversity 90° spatial optical mixer [110]

Crystal Optical Mixers

Material Properties of Crystal Optical Mixers

Crystal type optical mixers use the optical properties of uniaxial crystals, such as birefringence, double reflection, or electro-optical effects, to coherently mix the signal light and local oscillator light and generate a phase shift through the electro-optical effect. Commonly used crystal materials are lithium niobate, polymers, calcite, and yttrium orthovanadate [112].

Optical mixers require a low insertion loss, high isolation, high common mode rejection ratio, and small polarization loss. Therefore, crystal type optical mixers have high optical quality requirements for crystals:

1. The refractive index difference of crystals in the communication window band should be greater than 0.1.
2. The temperature sensitivity should be low, with stable performance in high- and low-temperature regions.
3. The transmittance in the window band should be high, and the insertion loss should be low.
4. They should have a small thermal expansion coefficient, small thermo-optical coefficient, and the suitable hardness should stay consistent.
5. Crystals are easy to grow, easy to use for optical processing, and inexpensive.

Table 29.1 lists the main birefringent crystal materials currently used in optical mixers and compares the parameter characteristics of the materials [113].

The yttrium orthovanadate (YVO_4) crystal has large refractive and discrete angles, excellent temperature stability and refractive index characteristics, and good physical and mechanical properties. The titanium dioxide (TiO_2) crystal has the largest refractive index difference; however, its extreme hardness makes optical processing difficult. Although lithium niobate ($LiNbO_3$) has mechanical and physical properties similar to YVO_4, its crystal growth technology is relatively mature. Its birefringence index is three times smaller than that of YVO_4, which increases the required crystal size. Therefore, YVO_4 is currently the best material for use in photoisolator birefringent crystals.

Uniaxial crystals are the main components for making optical mixers. The uniaxial crystal materials commonly used in making spectroscopic devices include calcite ($CaCO_3$), yttrium orthovanadate, and lithium niobate [114]. The specific properties and production and processing of these materials are compared, as shown in Table 29.2.

A synthetic positive uniaxial yttrium orthovanadate crystal has high transmittance, easy growth, moderate hardness, mature coating processing technology, easy

Table 29.1 Crystalline materials for optical mixers [113]

Crystal	YVO_4	TiO_2	$CaCO_3$	$LiNbO_3$
Space group	*I*41 /*amd*	*P*42 /*mnm*	$R\bar{3}C$	*R*3*C*
Birefringence (ne-no)	0.2039 (1.55 ìm)	0.256 (1.53 ìm)	−0.1572 (1.497 ìm)	−0.0738 (1.44 ìm)
Discrete angle (45°)	5.67°	5.63°	4.33°	1.96°
Through the band(μm)	0.4–5	0.4–5	0.35–2.3	0.4–5
Coefficient of thermal expansion $\times 10^{-6}$/°C	*c* axis: 11.4 *a* axis: 4.4	9.2 7.1	30.3 5.4	19.7 7
Mohs hardness	5	6.5	3	5
Deliquescence	None	None	Weak	None

Table 29.2 Comparison of the performance and production processes of crystals commonly used to make light beam separators

Crystal material	Optical quality	Transmittance	Birefringence	Processing hardness	Mass production
Yttrium orthovanadate	Good	Good	Very good	Moderate	OK
Lithium niobate	Good	Good	Bad	Moderate	OK
Calcite	Bad	Good	Good	Too soft	Difficult
Rutile	Good	Good	Very good	Too hard	Very difficult

mass production [115], and excellent birefringence performance [116]. The birefringence index difference is three times that of a lithium niobate crystal, that is, the beam of light has a large deviation angle, which can completely separate the beam of light in a short distance. This further improves the optical mixer integration [117, 119].

Crystal Optical Mixers

In 2003, Arkady et al. [118] used lithium niobate crystals to fabricate a 90° optical mixer. The structure of the optical mixer is shown in Fig. 29.19. The mixer has an in-phase optical path and an orthogonal optical path. The phase shifter and the Mach–Zehnder interferometer modulate the amplitude and phase of the two light channels. They control the voltage of the upper and lower arms of the four couplers to evenly distribute them at the output of the directional coupler; the phase difference between the two arms of the waveguide is 90°. The biggest advantage of the lithium niobate crystal chip is the high degree of internal integration, which greatly reduces the size of the entire optical mixer and enables the optical mixer to be integrated on a balanced detection chip (Fig. 29.17).

In 2009, Wan et al. [119] proposed an electronically controlled phase shifting spatial mixer based on crystal birefringence and electro-optical effects. The structure used the birefringence effect of a $LiNbO_3$ crystal to mix the local oscillator and signal lights and then used the electro-optical effect to shift the phase. Its basic structure is shown in Fig. 29.18. The size of the entire mixer is 50 mm × 20 mm × 20 mm, the adjacent size of the four output beams is 4.4 mm, and the control voltage is 100 V.

In 2010, Hou et al. [120] proposed, made, and experimentally verified an electronically controlled phase shift 2 × 4 90° crystal spatial light mixer. The specific principle involved splitting and coupling the signal light and local oscillator light through four crystals with the same structure, and the electronically controlled phase shift plate provided the required phase shift. The mixer had a controllable phase shift, but needed to provide power to the system.

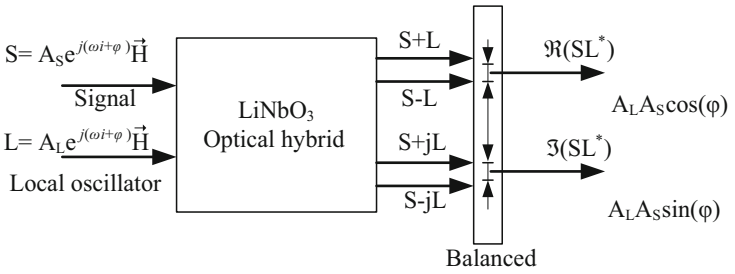


Fig. 29.17 Lithium niobate optical mixer [118]

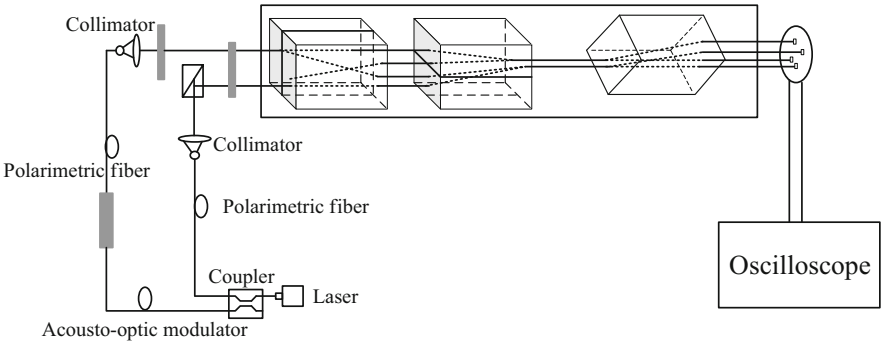


Fig. 29.18 90° spatial light mixer based on a birefringent crystal [119]

In 2013, Ke et al. [121] designed a 90° optical mixer based on crystal birefringence. To address the loose structure and poor temperature stability of discrete optical mixers, polarizing beam splitters, waveplates, yttrium orthovanadate crystals, and other devices, a spatial optical mixer with a compact structure, easy processing, and temperature insensitivity was designed.

In 2015, Zheng et al. [122] used two fast birefringent crystals, different phase-difference waveplates, and other optical components to design and then improve a tiled optical mixer. The improved optical path can be folded, reducing the size of the optical bridge in the mixer. However, the large number of structural elements in this design led to a large loss of light energy in the beam, resulting in a half-wave loss.

In 2018, Song et al. [123] designed a 2 × 4 90° spatial light mixer using electronically controlled birefringent crystals and polarizing birefringent crystals. Through the electronically controlled birefringent crystals, the signal light and local oscillator light are mixed by being double-reflected four times inside the crystal. The mixer could control the crystal voltage and the phase. It was simple in construction and easy to integrate.

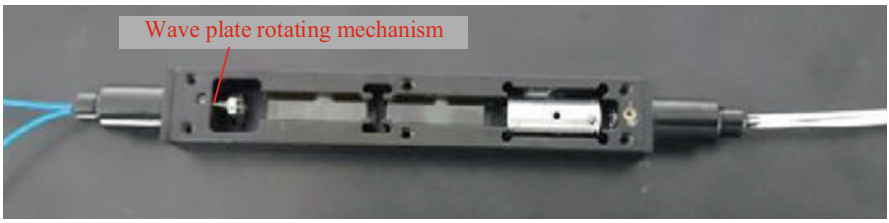


Fig. 29.19 2 × 4 90° crystal-shaped spatial light mixer [124]

In 2020, Han of Xi'an University of Technology [124] designed a 2 × 4 crystal-shaped 90° spatial light mixer based on the crystal birefringence effect, as shown in Fig. 29.21. Reference [124] describes the basic working principle and parameters. The integrated optical fiber and collimating lens at both ends of the mixer improved the quality of the beam entering the optical mixer. The waveplate was controlled by a rotating mechanism. The return loss of the test was 55 dB, and the external dimensions were 99 mm × 15 mm × 15 mm (Fig. 29.19).

The development of materials has produced optical mixers based on crystals. This type of optical mixer has the advantages of a low insertion loss, simple structure, and easy installation and adjustment, which are conducive to miniaturization and integration. However, because the crystal mixers control the phase through electro-optical effects, the crystal must be plated with electrodes; the device process is complex and additional power is required. Table 29.2 lists the properties of several commonly used crystals.

Other Types of Optical Mixers

In addition to the 90° optical mixer, the most commonly used mixers in coherent optical communication systems include 45°, 60°, and 120° optical mixers. Among them, the 120° optical mixer is essentially a 3 × 3 symmetric optical coupler. The coherent receiver based on the 120° optical mixer does not require a balanced detection module, which reduces the overall structure of the coherent receiver while saving costs [125].

In 1989, Nicholson et al. [126] first proposed a 120° optical mixer, and its basic structure is shown in Fig. 29.23. Based on the analysis, simulation, and demonstration of its basic principle, a 120° optical mixer was fabricated and used in a differential phase shift keying coherent optical communication system. When the laser frequency deviation was less than 1% of the bit rate, the performance of the coherent receiver based on the 120° optical mixer was consistent with that of a heterodyne receiver, and even more stable (Fig. 29.20).

In 2014, Dong et al. [127] designed a 120° coherent optical communication optical mixer based on a polarization prism. The mixer used an inverted-cone optical coupler integrated with the fiber and the chip, a polarization splitting prism, and a 3 × 3 MMI 120° optical mixer structure. The heterodyne communication

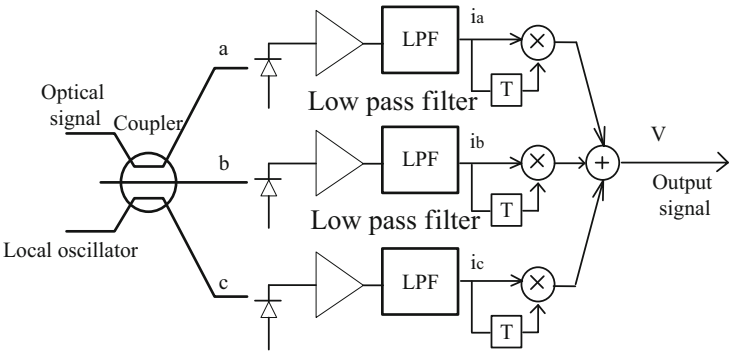


Fig. 29.20 Schematic diagram of a 120° optical mixer [126]

performance of a coherent optical receiver made of this type of optical mixer in the 1500-nm band can meet the standard of commercial coherent receivers.

Multimode Interference Coupler Optical Mixer

At present, domestic and foreign solutions for realizing optical mixers on planar waveguides mainly include star couplers [128], arrayed waveguide gratings [129], and multimode interference (MMI) couplers [130, 131]. Star couplers and arrayed waveguide grating optical mixers have complex structure designs, large device sizes, and poor transmission efficiency. Therefore, domestic and foreign research groups use multimode interference couplers for optical mixing.

According to the different implementations of MMI couplers, 90° optical mixers based on MMI couplers can be summarized into three categories: 4 × 4 MMI couplers [132], 2 × 2 MMI coupler combinations [133], and 2 × 4 MMI coupler combinations [134].

4 × 4 MMI Coupler Combination Types

In 2009, a 4 × 4 MMI coupler type 90° optical mixer was reported, based on the InP platform [124, 129]. The device size was 15 × 497.5 mm^2, and the additional loss in the C band range was less than 1 dB.

In 2011, Voigt et al. [135] reported a 90° optical mixer for silicon nanowire waveguides fabricated on a silicon-on-insulator (SOI) platform. The device has a size of 10 × 185 mm^2, an additional loss of less than 0.5 dB, and a phase deviation of less than 5° in the C band range. Compared with the 90° optical mixer made of four SOI ridge waveguides, a better phase accuracy was obtained.

In 2011, Fujitsu Corporation [136] proposed a wedge-shaped 4 × 4 MMI coupler type 90° optical mixer on the InP platform. The optical mixer had an additional loss of less than 1.4 dB in the 1530 nm–1600 nm wavelength range, and the phase deviation was less than ±5.4°.

In 2018, Xu et al. [137] reported a 90° optical mixer designed on the SOI platform, which utilized a subwavelength grating to adjust the effective refractive index of different modes in the waveguide. It then changed the beat frequency to achieve an efficient self-mapping effect.

In 2018, Wang et al. [138] first proposed a 90° optical mixer based on a hybrid plasmon waveguide, as shown in Fig. 29.13. The phase deviation at the 1550 nm wavelength was less than 0.7°. However, its additional loss was large and it was only suitable for short-distance coherent optical systems.

In addition to 90° optical mixers that can be implemented on platforms with high relative refractive indices, such as InP and SOI, optical mixers can also be implemented on polymer platforms with low relative refractive indices [139]. These devices tend to be low cost and have good polarization demultiplexing performance; however, owing to polymer limitations, the bandwidth and durability of the device are often limited. There are also reports of 4 × 4 MMI coupler-based optical mixers using silicon nitride waveguides with moderate refractive indices [140].

2 × 2 MMI Coupler Combination Types

In 2010, Japan's NTT Corporation [141] reported a 90° optical mixer with a refractive index difference of 1.5% on a silicon dioxide platform. The 90° optical mixer consisted of four 2 × 2 MMI couplers plus a 90° phase shifter. In the C band and L band ranges, the 90° optical mixer had a phase deviation of less than 0.7°.

In 2011, NTT Corporation [142] proposed replacing one of the traditional four 2 × 2 MMI couplers with a Y branch, which reduced the use of a 90° phase shifter. In the C band and L band ranges, the 90° optical mixer phase deviation was less than 3°.

In 2017, Elenion Technologies [143] reported a low-loss 90° optical mixer on the SOI platform. The optical mixer used a novel topology including a Y branch, three 2 × 2 MMI couplers, and four 90° curved waveguides. The size of the entire device was reduced to 21 mm × 31.9 mm. In the C band range, the insertion loss of the device was less than 0.5 dB, the phase deviation was less than 30°, and the common mode rejection ratio was less than 30 dB.

Different Phase Demodulation Types of MMI Couplers

To obtain higher spectral efficiency, the *m*-ary phase modulation format is used. In addition to common optical mixers are optical mixers with different phase demodulations.

In 2010, Fujitsu [144] reported a 45° optical mixer based on the InP platform, which was based on a combination of a 2 × 8 MMI coupler, three pairs of phase shifters, and three 2 × 2 MMI couplers. In a 32-nm bandwidth, the phase deviation was less than 5°. This optical mixer was smaller in size than traditional octal differential phase shift keying (DPSK) demodulators.

In 2011, Fujitsu Corporation [145] proposed a 60° optical mixer based on the InP platform. It was based on a combination of a 2 × 6 MMI coupler, two pairs of phase

shifters, and two 2 × 2 MMI couplers and was integrated with a balanced monolithic photodetector. At a 40-nm bandwidth, the phase deviation of the device was less than 5°.

In 2017, a Canadian group [146] reported a 120° optical mixer based on a 3 × 3 MMI coupler. The device had a phase deviation of less than 10° in the 55-nm wavelength range of the C band and L band and could be applied to high-speed passive optical networks (PONs). In 2021, the Institute of Semiconductors, Chinese Academy of Sciences [147], designed a 72° optical mixer based on a non-overlapping mapping of 2 × 5 MMI couplers on the SOI platform. The device length was 89.5 mm and the width was 8 mm.

MMI Optical Mixer Classification

Owing to the different positions of the input light field, the light-field modes excited by multimode waveguides are different. According to the position of the input light field and the excited light-field mode, MMI couplers are classified into general interference, paired interference, and symmetric interference MMI couplers [148].

Interference-Coupled MMI Coupler

All light field modes in the interferometric MMI coupler are excited; hence, the position of the input light field does not affect the mixing efficiency or other performance parameters of the optical mixer. When the multimode waveguide length $L = p(3L_\pi)$, $p = 0, 1, 2, \ldots$, the output optical field expression in the multimode waveguide is

$$\psi(x, L) = \sum_{v=0}^{m-1} C_v \psi_v(x) \cdot (\pm 1). \tag{29.14}$$

In the formula, C_v is expressed as the field excitation coefficient of the v-order mode, $\psi(x, L)$ is the output light-field mode, and $\psi_v(x)$ is the mode field distribution of the v-order mode. At this time, the light field of the output port is distributed.

The reproduction or mapping of the light field of the input port can obtain the length of N reimages as

$$L = \frac{P}{N}(3L_\pi), \tag{29.15}$$

where L is the length of the modular field and N is the multiplicity of the mapping. There is no common divisor between P and N. For a more compact design, the MMI coupler selects the first self-mapping point. At this time, $P = 1$. Using Fourier space decomposition, the form of the output light field after the input light field passes through the length L can be obtained:

$$\psi(x,L) = \sum_{v=0}^{\infty} C_v \psi_v(x) \exp(-j\beta_v L), \tag{29.16}$$

where C_v is expressed as the field excitation coefficient of the v-order mode, $\psi(x, L)$ is the output light-field mode, and $\psi_v(x)$ is the mode field distribution of the v-order mode.

Paired-Interferometric MMI Coupler

The paired-interferometric MMI coupler can only excite the light-field mode when the pilot-mode order is divisible by three. Its excitation coefficients are zero for all other excitation modes. Therefore, the position of the input light field will affect the light field excitation. The self-map of the input port can be obtained at the output port. At this time, the length of N-fold self-influence is

$$L = \frac{p}{N} L_\pi, \tag{29.17}$$

where L is the length of the mode field and N is the multiplicity of the map. There is no greatest common divisor between P and N; $P = 1$. When the width of the MMI coupler is the same, the length of the paired interferometric coupler is only half that of the interferometric coupler. Therefore, this principle can be used to reduce the length of the MMI coupler [149].

Symmetrical Interferometric MMI Coupler

In the multi-channel waveguide of a symmetric-interferometric MMI coupler, only even-order modes can be excited; the excitation coefficient of the odd-order modes is 0. Therefore, the position of the input light field must be located in the center of the waveguide. When the length of the multimode waveguide $L = P(3L\pi/4)$, the output port obtains the reproduction or mirror of the input square. At this time, the length of the N-fold mirror is

$$L = \frac{P}{N} \frac{3L_\pi}{4}, \tag{29.18}$$

where L is the length of the mode field and N is the multiplicity of the image. There is no common divisor between P and N; $P = 1$. When the width of the MMI coupler is the same, the length of the symmetrical interferometric MMI coupler is only one-fourth that of the interferometric MMI coupler. The phase relationship of the output port of the MMI coupler is

$$\varphi_{\mathrm{rs}} = \frac{\pi}{4N}(S-1)(N-S). \tag{29.19}$$

Through the above analysis, we can summarize the imaging rules of MMI couplers of different interference types, as shown in Table 29.3 [149].

Table 29.3 Summary of MMI couplers with different interference mechanisms

Interference mechanism	General interference	Pairing interference	Symmetric interference
Input waveguide × output waveguide	$N \times N$	$2 \times N$	$1 \times N$
Input location	Any location	$y = \pm \frac{w_e}{6}$	$y = 0$
First image location	$3L_\pi$	L_π	$(3L_\pi)/4$
First N remapping position	$3L_\pi/N$	L_π/N	$(3L_\pi)/4N$
Mode-incentive condition	None	$C_v = 0$, $v = 2, 5, 8, \ldots$	$C_v = 0$, $v = 1, 3, 5, \ldots$
Multimode-zone length	$L = \frac{p}{N} 3L_\pi$	$L = \frac{p}{N} L_\pi$	$L = \frac{p}{N} \frac{3L_\pi}{4}$

Arrayed Waveguide Grating Structured Optical Mixer

An arrayed waveguide grating (AWG) is a planar optical mixing device based on the interference and diffraction effects of spatial light. It is fabricated using planar lightwave circuit (PLC) technology, and the process is compatible with the complementary metal oxide semiconductor (CMOS) process [150].

Compared with other wavelength division multiplexing devices, the AWG fabrication process is simple, the number of channels is large, the insertion loss is low, the device size is small, and it can easily be coupled with optical fiber. In addition, an AWG can be integrated on-chip with other integrated photonic devices. AWGs are not only used in the communication field [151], but are also widely used in optical-fiber sensing [152].

Basic Principle of the Arrayed Waveguide Grating

The basic structure of an arrayed waveguide grating (AWG) includes an input/output waveguide, two focusing-plate regions, and a multi-channel waveguide phased array. ΔL is the constant physical path length difference between adjacent waveguides of the phased array. The working process of the AWG can be briefly described as follows. When the signal light carrying information enters the input star coupler flat-plate waveguide through the central input waveguide, it disperses and diffracts and then reaches the end face of the arrayed waveguide with the same phase and is coupled to the arrayed waveguide region. This region is collected by each individual waveguide and propagates independently. Because the end faces of the arrayed waveguides are located on the circumference of the grating circle, the diffracted light is coupled to arrayed waveguides with the same phase.

Since the adjacent waveguides in the arrayed waveguide have the same length difference ΔL, the diffracted light will have an equal phase difference after passing through the arrayed waveguide. A mirror image of the light field in the input star coupler flat waveguide is generated at the output star coupler flat waveguide. Because the optical paths of different wavelengths propagating through the arrayed waveguide cause different phase differences, different light wavelengths will be output to and from different channel positions, respectively [153].

Research Progress on Arrayed Waveguide Gratings

Donald et al. [154] proposed the concept of combining planar waveguides and planar gratings in 1987, and Smith proposed the structure of arrayed waveguide gratings in 1988; it is regarded as the prototype of the AWG structure [155].

In 2010, Vellekoop and Smith [156] realized the first arrayed waveguide grating (AWG) element for use at multiple wavelengths. In 2012, Zirngibl et al. first proposed an InP-based element, which used a smaller waveguide size and bending radius in the design, and could be well integrated with other components [157].

In 2000, McLaughlin et al. proposed a design for machining the input and output ends of the AWG flat waveguide part into a tapered port. This design made the transmission spectrum more planar, reduced the coupling loss of the device, and effectively controlled the temperature and polarization [149]. In the same year, an AWG was officially commercialized in a dense wavelength division multiplexing system and played a relatively good role [158].

In 2005, Hill et al. designed a 4 × 4 AWG with a 1550 nm center wavelength and 2 nm channel spacing [159]. In 2009, Heck et al. achieved eight channels, channel spacing of 0.84 nm, and a 3 dB bandwidth of 0.8 nm [152].

In 2017, Stanton et al. [160] proposed and prepared an arrayed waveguide grating with an eight-channel silicon-on-insulator, with a 3.7 mm^2 device size, a 0.08 dB–0.81 dB channel loss, and a −30.2 dB channel crosstalk, which had better performance.

In 2017, Fang et al. designed and prepared a 1 × 4 cascaded AWG, which used a cascaded method to suppress crosstalk between channels. Compared with a single AWG, the cascaded AWG crosstalk increased to −33.2 dB, and the device size was 1.0 mm × 1.4 mm [161]. In 2017, Song and He et al. designed and fabricated a 3.2 nm channel spaced, 16-channel silicon-based AWG with a 0.58 mm × 0.77 mm device size and an insertion loss of 4.9 dB [162]. Cascading the devices is one solution to give the AWG a greater bandwidth and higher resolution.

In 2018, Stanton et al. [163] achieved a low-loss arrayed waveguide grating at a 2 μm wavelength, with a channel loss of 2.4 dB, and −31 dB crosstalk per channel. In the same year, Elke et al. [164] fabricated an AWG with an 850 nm center wavelength using a low-cost, rapid prototyping, maskless laser direct writing method.

In 2021, Park et al. fabricated an eight-channel AWG with an 838-nm center wavelength based on a silicon nitride platform. After testing, the insertion loss of this device was 0.43 dB and the inter-channel crosstalk was −29.6 dB [165].

Owing to the relatively low refractive index of SOI using silicon nitride materials, the device requires a larger area. For this reason, in 2020, Menard et al. proposed the concept of a waveguide superlattice, which reduced the size of AWGs prepared based on silicon nitride materials more than twice, while achieving an insertion loss of 3.4 dB and a channel crosstalk of −18 dB [159].

In 2021, Liu et al. designed a four-way silicon-based AWG device with an insertion loss of 2.5 dB [160]. In 2020, Hon et al. of the City University of Hong Kong successfully fabricated an AWG device with an 860-nm center wavelength and

40 channels on a silicon nitride platform. The device had a bandwidth of more than 60 nm and a center-channel insertion loss of 1.3 dB [166]. This work proved the feasibility of using large-scale programmable logic controller (PLC) technology to fabricate arrayed waveguide gratings, which further improved the degree of integration.

In 2021, Chen successfully fabricated 15 channels using silicon nitride on an SOI wafer and reduced the device size by combining it with a silicon photonic-loop reflector. The actual measured insertion loss was 5.46 dB, and the channel crosstalk was 19.2 dB [167]. In 2021, Taha et al. successfully developed a compact four-channel AWG device on a silicon photonic platform based on MMI, with a device size of 367 × 67 square microns and an insertion loss of 9.6 dB [168].

Comparison of the Material Properties of Optoelectronic Devices

Common Materials for Optical Mixers

In the optical communication field, the integration and flattening of coherent optical communication systems has promoted the integrated development of optoelectronic devices. Optical receiving and detector devices, including optical mixers, are oriented towards a tunable, focused response and narrow linewidth. The transmission linearity of the optical waveguide in the dielectric layer laid the foundation for integrated optics, after which optical waveguide–based optical mixer structures have been continuously produced.

Many materials with good optical properties can be used to make optical waveguides, among which five are the most widely used and more mature technologies [169]: optical waveguides based on lithium niobate crystals, silicon dioxide waveguides, semiconductor optical waveguides, optical glass waveguides, and polymer waveguides.

In the integrated optic field, the lithium niobate crystal ($LiNbO_3$) has excellent physical properties, such as electro-optics, acousto-optics, nonlinearity, photorefractivity, photoelasticity, and laser activity [170]. The lithium niobate crystal has a cubic crystal structure. One of its advantages is that it has a large electro-optical coefficient, and its optical-switching voltage is relatively low. Its outstanding electro-optic properties make it widely used for electro-optic devices, such as optical phase or intensity modulators, optical phase retarders, optical mixers, and optical scanners [171].

The silicon dioxide optical waveguide has the advantages of matching the production process of other electronic circuit fabrication processes, stable physical and chemical properties, good compatibility with optical fibers (high coupling efficiency), and extremely small loss in the strip-shaped single-mode waveguide. The upper and lower cladding layers of the silicon dioxide optical waveguide are both silicon dioxide, and the core layer is Ge-doped silicon dioxide. The preparation usually adopts lithography, reactive ion etching (RIE), and annealing processes.

Ion exchange is the most commonly used glass-manufacturing process. During the manufacturing process, cations (usually Na+ ions) in the glass material are replaced by ions with larger sizes or higher polarizability, making the refractive index of the ion-replacement region higher than that of other regions to form a waveguide structure. The ion exchange process is simple and low cost [172].

Polymer optical waveguides have been a research hotspot in recent years. They have the advantages of a simple fabrication process, low cost, wide bandwidth, and high photoelectric-coupling efficiency. They are suitable for fabricating passive waveguide devices. However, they have poor thermal stability and poor mechanical properties.

Comparison of Material Properties

The phase deviation of waveguide-type optical mixers is relatively large, so higher-order optical mode interference is used to reduce the phase deviation. The high-order mode of generated light places high fabrication requirements on the substrate material, and the fabrication technology is complicated. In addition, there are unavoidable additional signal losses in multiple input and output MMI couplers, as well as problems, such as signal crosstalk between multiple outputs. In addition, the performance of waveguide mixers of different materials is very different, so we compare different waveguide materials, as shown in Table 29.4.

Mixing-Error Analysis

The use of discrete optical elements, such as waveplates, polarization splitting prisms, and collimating lens groups, in the optical mixer, as well as the refractive

Table 29.4 Performance comparison of different waveguide materials

Material	Transmission loss (1550 nm, dB/cm)	Coupling loss (dB/ end face)	Core /core relative refractive index difference	Advantages and disadvantages
Lithium niobate	0.5	1	0–0.5%	Large size and small refractive index difference; belongs to weak refractive index optical waveguides
Silica	0.05	0.25	0–4%	Low transmission loss, good refractive index matching, low coupling loss
Polymer	0.1	0.1	0–35%	Devices are susceptible to contamination; aging leads to unstable performance
Silicon	0.05	0.1	0–0.5%	Low-loss, complex fabrication process, uneven refractive index distribution, unstable device performance

index distribution and assembly accuracy of the optical-fiber structured optical mixer will have a greater impact on the performance of the optical mixer. In addition, the vibration, temperature will also have a great impact on the stability of the device; thus, various optical mixer errors are theoretically analyzed [173].

Phase Error

A large number of optical elements are used in an optical mixer. The thermal expansion of an element caused by temperature, assembly errors, and aberrations in the optical element itself will cause the optical path difference of the passing light to not be strictly equal. It will increase with the increase of the laser wavelength and gradually expand [174].

The phase difference, the most important performance index of an optical mixer, is affected by many factors, such as the deviation of the optical axis of the waveplate, the phase delay deviation of the waveplate, and the phase delay deviation of the polarization splitter. According to the definition and calculation formula of the optical path difference, the additional phase difference is proportional to the frequency difference between the signal light and the local oscillator and the difference between the lengths of the two arms. When the frequency difference is 30 GHz and the difference between the lengths of the two arms is 1 mm, the additional phase difference is 0.628 mrad. When the frequency difference is below GHz, the additional phase difference can be ignored.

Polarizing-spectroscopic film is a multilayer film that alternately stacks high and low refractive index films on the spectroscopic surface so that the incident ray meets the Brewster angle, the P light transmittance is close to 100%, and the S light reflectivity is close to 100%. In practice, the phase effect of the polarizing spectroscopic film cannot reach the ideal state, so the corresponding differences of 90° or 180° cannot be accurately satisfied, which is also one of the important factors affecting the relative phase difference.

Insertion Loss

Each optical element in the optical mixer is a discrete element. When the laser is incident on and off the element surface, there will inevitably be insertion loss, which will reduce the sensitivity of the optical receiver where the optical mixer is located. The following types of optical components in the optical mixer are the main factors causing insertion loss.

Fiber Collimating Lens

As shown in Fig. 29.19, the signal light and the local oscillator light are respectively input to the optical element at the rear end, after being output from the pigtail of the laser by a fiber collimating lens to collimate the light beam. The four-way output light at 0°, 90°, 180°, and 270° is also coupled into the fiber through the collimating lens, and the light intensity of the output light is detected by the PIN. Because the lens focus in the collimating mirror is located at the fiber core with the highest coupling efficiency, and the diameter of the single-mode polarization-maintaining

fiber core is less than 10 mm, the coupling efficiency of the spatial light beam and fiber-collimating lens is low. Thus, the collimating lens at 90° is the primary factor for optical mixer insertion loss.

Polarization Beam-Splitting Prism (PBS) and Waveplate

Multiple polarization-splitting prisms are used in optical mixers. Their basic function is to separate the polarization components of the transmitted P ray and the reflected S ray. For the P and S lights, the transmittance of the prism is $(1 - R)^2$. The thickness of the waveplate is only about 1 mm, and the loss of optical power can be ignored. In addition, the surface is coated with the same anti-reflection film as the polarization splitting prism. Therefore, in theory, the transmittance of the waveplate is $(1 - R)^2$.

Splitting Ratio Error

The splitting ratio error refers to the power ratio between two adjacent channels. Theoretically, their ratio should be 1:1. However, since $T_P \neq R_S$ of the polarization-splitting prism, and the direction of the fast (slow) axis of the $\lambda/2$ waveplate is not strictly equal to 29.5°, the powers of the P light and S light of the outgoing polarizing-splitting prism are unequal. This means that the DC components of the adjacent two optical currents with a 180° difference are also not strictly equal, resulting in the corresponding noise.

Although the direction of the fast (slow) axis of the $\lambda/2$ waveplate cannot be accurately adjusted to 29.5° by the scale on the adjustment rack, the corresponding angular error can actually be compensated by rotating the $\lambda/2$ waveplate.

The polarization splitting prism introduces an inherent splitting ratio error. In the two adjacent channels obtained from the polarization-splitting prism data, the ratio of the reflected light to projected light power should be $R_S : T_P \approx 1.041$.

Analysis of the Mixing Efficiency of an Optical mixer in a Coherent Optical Communication System

Light Field Pattern Matching

In coherent optical communication, matching the light-field modes refers to matching the amplitude distribution of the light field between the signal light and the local oscillator light. This determines the energy distribution of the light spot received by the mixed light on the detector target surface, and the coincidence and uniformity of the light spot will inevitably affect the mixing efficiency of the coherent detection system [170]. Therefore, amplitude distribution matching is an important coherence premise and is very important for whether the signal and local oscillator lights can achieve high-efficiency mixing.

Optical Field Distribution

In coherent optical communication systems, the signal light and local oscillator light usually have several types of light-field distributions. The laser light-field distribution can be divided into Airy, uniform, and Gaussian distributions [170]. The different light-field distribution types affect the mixing efficiency of the mixer [171]. Several types of light-field distributions are as follows:

Gaussian distribution (Gauss):

$$U(r) = \exp\left(\frac{-r^2}{\omega_0^2}\right). \tag{29.20}$$

Airy distribution (Airy):

$$U(r) = \frac{2J_1\left(\frac{\pi r}{F\lambda}\right)}{\frac{\pi r}{F\lambda}}. \tag{29.21}$$

Uniform distribution:

$$U(r) = 1. \tag{29.22}$$

J_1 is a first-order Bessel function, F is the transmission parameter of the optical system, which usually determines the size of the light spot of the Airy spot; λ is the wavelength of the laser, and ω_0 represents the waist radius of the Gaussian light beam. The optical signal can have a Gaussian distribution after transmission, and an approximately uniform distribution after a long-distance transmission. Therefore, we describe the optical mixer efficiency when the signal light-field distribution is Gaussian and uniform, and the local oscillator light is three cases.

Mixing Efficiency

According to the coherent detection principle and limited by the bandwidth of the photodetector, a beat signal is obtained by mixing the wavefronts of the signal and the local oscillator light. The DC component is removed after subsequent processing to obtain an intermediate frequency signal with modulation information. The signal can be described in the following form:

$$i_{\mathrm{AC}} = \frac{R}{2Z}\iint_D A_{\mathrm{S}}(r,\varphi)A_{\mathrm{LO}}(r,\varphi)\cos(\omega_{\mathrm{d}}t,\Delta\varphi)ds, \tag{29.23}$$

where R represents the photodetector response, Z represents the impedance characteristics of the photodetector, and D represents the receiving photosensitive area of the photodetector. $A_{\mathrm{s}}(r, \varphi)$ and $A_{\mathrm{LO}}(r, \varphi)$ represent the amplitude distribution functions of the signal light and local oscillator light on the receiving plane, respectively, where r represents the off-axis distance. In the coherent detection process, the shot noise caused by the local oscillator light is the main source of the mixer noise.

The DC signal corresponding to the local oscillator light is used to represent the DC signal output in the detector, and the photocurrent can be written as shown in Eq. (29.24).The signal-to-noise ratio of the receiving system is the intermediate frequency signal photocurrent and the DC component mean-square ratio generated by the shot noise of the local oscillator light as Eq. (29.25). p_s indicates the power of the signal light, B indicates the bandwidth of the shot noise caused by the local oscillator, and η indicates the mixing efficiency of two beams of light.

$$i_{\mathrm{DC}} = \frac{R}{2Z}\iint_D U_{\mathrm{LO}}^2(r,\varphi)dS \tag{29.24}$$

In the formula, R represents the photodetector response, Z represents the impedance characteristics of the photodetector, and D represents the receiving photosensitive area of the photodetector:

$$\mathrm{SNR} = \frac{\langle i^2\mathrm{AC}\rangle}{2qBi_{\mathrm{DC}}} = \eta\frac{RP_{\mathrm{S}}}{qB}. \tag{29.25}$$

The mixing-efficiency formula can be established as follows:

$$\eta = \frac{\left|\int_0^{2\pi}\int_0^{r_0} U_{\mathrm{S}}(r,\varphi)U_{\mathrm{LO}}(r,\varphi)\exp[i\Delta\psi(r,\varphi)rdrd\varphi]\right|^2}{\int_0^{2\pi}\int_0^{r_0}[U_{\mathrm{LO}}(r,\varphi)]^2 rdrd\varphi\int_0^{2\pi}\int_0^{\infty}[U_{\mathrm{S}}(r,\varphi)]^2 rdrd\varphi}, \tag{29.26}$$

where $U_{\mathrm{S}}(r, \varphi)$ and $U_{\mathrm{LO}}(r, \varphi)$ are the normalized amplitude distributions of the signal light and local oscillator light, respectively, and the integral interval is the area of the entire photosensitive surface.

When the local oscillator light is uniformly distributed, its amplitude distribution is $U_{\mathrm{LO}}(r) = 1$, and the light field distribution function of the signal light and the local oscillator light is substituted into the above mixing-efficiency formula, and the common value of the mixing efficiency changing with the detector is obtained:

$$\eta(r_0) = 4\frac{\left[1 - J_0\left(\frac{\pi r_0}{F\lambda}\right)\right]}{\left(\frac{\pi r_0}{F\lambda}\right)^2}. \tag{29.27}$$

The above formula is the mixing efficiency formula when the local oscillator light is uniformly distributed. In the formula, J_1 is the first-order Bessel function, F is the transmission parameter of the optical system, and λ is the wavelength of the laser.

When the intrinsic light has an Airy distribution, the light-field distribution of the signal light and the local oscillator light is substituted into the mixing-efficiency

formula to obtain the mixing-efficiency formula when the local oscillator light has an Airy distribution:

$$\eta_{\mathrm{AA}}(r_0) = 1 - J_0^2\left(\frac{\pi r_0}{F\lambda}\right) - J_1^2\left(\frac{\pi r_0}{F\lambda}\right). \tag{29.28}$$

When the local oscillator light has a Gaussian distribution, the optical field amplitude distribution function of the two lights is also substituted into the mixing efficiency Eq. (29.26), sorted out as above, and the set parameters are substituted. When the local oscillator light has a Gaussian distribution, the mixing efficiency formula changes with the detector radius r_0:

$$\eta_{\mathrm{AG}}(r_0) = \frac{8\left|\int_0^{r_0} J_1\left(\frac{\pi r}{F\lambda}\right).\exp\left(-\frac{r^2}{\omega_0^2}\right)dr\right|^2}{\omega_0^2\left[1 - \exp\left(-2\frac{r_0^2}{\omega_0^2}\right)\right]}. \tag{29.29}$$

Light-Spot Size Mismatch

In a coherent laser communication system, when the amplitude and phase of the signal light and the local oscillator light completely match, the coherence is the best and the detection sensitivity is the highest. In this case, the ratio between the size of the signal light and the local oscillator light spot and the half-pass of the detector is smaller. Usually, the receiving target surface of the optical mixer is placed at the waist of the laser beam, which is more conducive to concentrating the available energy, which, in turn, helps to achieve higher efficiency mixing.

In this case, the mixing efficiency formula of the optical mixer under different light spot sizes can be obtained:

$$\eta = 2\frac{\left[\frac{2}{k}\int_0^{X_0} J_1(X)J_0(2F\theta X)\exp\left(-\frac{X^2}{k^2}\right)dX\right]}{1 - \exp\left(-2\frac{X_0^2}{k^2}\right)}, \tag{29.30}$$

where $k = \frac{X}{Z} = \frac{\pi\omega}{F\lambda}$. $X = \frac{\pi r}{F\lambda}$, $Z = \frac{r}{\omega}$, and the integral area is the area where the signal light and the local oscillator spot coincide. When the signal light beam has a great influence on the mixing efficiency of the mixer, assuming that the entire local oscillator light is received at this time, the optical mixer efficiency formula can be expressed as

$$\eta = \frac{8}{k^2}\left[1 - \exp\left(-\frac{k^2}{4}\right)\right]^2. \tag{29.31}$$

Good light-spot size matching can increase the signal on the target surface of the detector, the effective coherent area of the local oscillator light, and the energy concentration, resulting in a higher mixing efficiency. By optimizing the detector size and the relative light beam size, the communication light beam and the system can be designed to improve the mixing efficiency, thereby improving the communication performance of the coherent laser communication system.

Wavefront Mass

The mixer adds gain to the detection sensitivity of the receiving system by coherently mixing the signal light with the wavefront of the local oscillator light. Phase matching, as a key coherence condition, is very important for achieving high-efficiency mixing. The wavefront quality of the light beam directly affects the phase change of the two lights in the coherence process, and the phase distortion of the optical wavefront has a non-negligible impact on the mixing efficiency. Two main elements affect the wavefront quality [175]:

1. Internal optical system: Atmospheric channels, temperature changes, shock-vibration stress, other natural factors, and platform vibration will affect the imaging quality of the optical system in the transmission orbit. For example, large temperature changes lead to changes in the refractive index and lens-surface curvature. Thickness changes, lens-spacing changes, etc., cause a phase distortion of the light wavefront.
2. Incident ray channel: The random variation of the atmospheric refractive index will produce atmospheric turbulence, which has a huge impact on the wavefront phase of the light beam. It not only causes a light intensity fluctuation of the light field on the detector target surface, but can also distort the light wavefront. This can cause a laser speckle effect, resulting in a decrease in the Strehl ratio of the light spot, and even a serious deformation in the amplitude distribution of the light field received by the light spot on the target-detector surface.

Zernike polynomials are typically used to describe the wavefront distortion phase of the signal light. The Zernike polynomial of the main quintuple primary phase can be written in the following form:

$$\begin{cases} \varphi_{\text{tilt}}(\rho,\theta) = a_{111}\rho\cos\theta \\ \varphi_{\text{defocussing}}(\rho,\theta) = a_{020}\rho^2 \\ \varphi_{\text{spherical aberration}}(\rho,\theta) = a_{040}\rho^4 \\ \varphi_{\text{coma}}(\rho,\theta) = a_{131}\rho^3\cos\theta \\ \varphi_{\text{astigmatism}}(\rho,\theta) = a_{222}\rho^2\cos^2\theta \end{cases} . \tag{29.32}$$

where ρ is used as the normalized radius, and a is used as the aberration coefficient to represent the size of the wavefront aberration. An equation for studying the effect of

wavefront distortion on the mixing efficiency of spatial light is established, and the mixing-efficiency formula is obtained as follows [177]:

$$\eta = \frac{\left| \int_0^{2\pi} \int_0^{r_0} E_S(r,\varphi) E_{LO}(r,\varphi) \exp[i\Delta\psi(r,\varphi) r dr d\varphi] \right|^2}{\int_0^{2\pi} \int_0^{r_0} [E_{LO}(r,\varphi)]^2 r dr d\varphi \int_0^{2\pi} \int_0^{\infty} [E_s(r,\varphi)]^2 r dr d\varphi}. \tag{29.33}$$

The above formula is the mixing-efficiency formula under the influence of different aberrations. Both the local oscillator light and the signal light have Gaussian distributions. The influence of the wavefront distortions caused by the five primary phase differences on the mixing efficiency of the optical mixer is analyzed. The Zernike polynomial expression of each aberration is added to the mixing-efficiency expression to obtain the mixing-efficiency expressions of five basic aberrations, including tilt, defocusing, spherical, coma, and astigmatism.

Equation (29.34) shows the mixing-efficiency expression under the influence of a tilt aberration:

$$\eta_{\text{Tilt}} = \frac{8\left| \int_0^{r_0} \exp\left(-\frac{r^2}{\omega_0^2}\right) J_0(k a_{111} \rho) r dr \right|^2}{r_0^2 \omega_0^2 \left[1 - \exp\left(-\frac{2r_0^2}{\omega_0^2}\right)\right]}, \tag{29.34}$$

where the integral area is the entire receiving photosensitive surface area, ρ is a normalized half, a is the Zernike aberration coefficient, ω_0 is the waist radius of the light beam, and r_0 is the detector radius.

Equation (29.35) is the expression of the mixing efficiency under the influence of a defocus aberration:

$$\eta_{\text{Defocus}} = \frac{8\left| \int_0^{r_0} \exp\left(-\frac{r^2}{\omega_0^2}\right) \exp(-ik a_{020} \rho^2) r dr \right|^2}{r_0^2 \omega_0^2 \left[1 - \exp\left(-\frac{2r_0^2}{\omega_0^2}\right)\right]} \tag{29.35}$$

Equation (29.36) is the expression of the mixing efficiency under the influence of a spherical aberration:

$$\eta_{\text{Spherical aberration}} = \frac{8\left| \int_0^{r_0} \exp\left(-\frac{r^2}{\omega_0^2}\right) \exp(-ik a_{040} \rho^4) r dr \right|^2}{r_0^2 \omega_0^2 \left[1 - \exp\left(-\frac{2r_0^2}{\omega_0^2}\right)\right]}. \tag{29.36}$$

Equation (29.37) is the expression of the mixing efficiency under the influence of a coma:

$$\eta_{\text{Coma aberration}} = \frac{8\left|\int_0^{r_0} \exp\left(-\frac{r^2}{\omega_0^2}\right) J_0(ka_{131}\rho^3) r dr\right|^2}{r_0^2\omega_0^2\left[1 - \exp\left(-\frac{2r_0^2}{\omega_0^2}\right)\right]} \tag{29.37}$$

Equation (29.38) is the expression of the mixing efficiency under the influence of astigmatism:

$$\eta_{\text{Astigmatism}} = \frac{8\left|\int_0^{r_0} \exp\left(-\frac{r^2}{\omega_0^2}\right) \exp\left(-ika_{222}\frac{\rho^2}{2}\right) J_0\left(ka_{222}\frac{\rho^2}{2}\right) r dr\right|^2}{r_0^2\omega_0^2\left[1 - \exp\left(-\frac{2r_0^2}{\omega_0^2}\right)\right]}. \tag{29.38}$$

A tilt aberration has the greatest impact on the mixing efficiency. It is mainly caused by the shaking of the sight axis, which needs to be compensated by improving the tracking and aiming accuracy. The influence of astigmatism is minimal; a 0.1 λ astigmatism reduces the mixing efficiency by 3%. Spherical and defocusing aberrations of 0.1 λ reduce the mixing efficiency by 5% and by 14% at 0.2 λ. The mixing efficiency can be maintained above 65% when the size of each aberration is within 0.2 [176].

Vibration and Tracking Errors

The optical axis direction of the signal light and local oscillator light in the actual transmission process is affected by angular deviation. For example, a larger-scale atmospheric turbulence causes the light beam to drift, so that the center position of the signal-beam propagation plane is offset.

The vibration of the external platform and the shaking of the optical axis caused by the loop noise of the acquisition, pointing, and tracking (APT) subsystem directly affect the coaxiality between the received two beams of light. It moves the focal plane of the light spot, which, in turn, affects the effective coherence area of the light beam and loses useful light information energy. The signal caused by the shaking of the optical axis and the offset of the local oscillator light field and the focal plane of the light spot can be represented as shown in Fig. 29.21.

The influence of the center distance between the signal light and the local oscillator light spot on the mixing efficiency is analyzed, and the effective coherent light-spot area and the mixing efficiency are written as

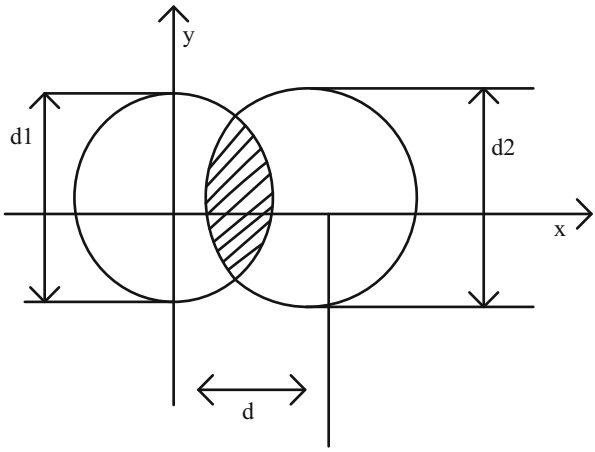


Fig. 29.21 Schematic diagram of the optical axis offset

$$s = \frac{\pi \bar{d}^2}{4} - \frac{\bar{d}^2}{2} \arcsin\left(\frac{d}{\bar{d}}\right) - \frac{\bar{d}^2}{4} \sin\left[2\arcsin\left(\frac{d}{\bar{d}}\right)\right] \tag{29.39}$$

$$\eta = \frac{S}{\pi(\bar{d}/2)^2} = 1 - \frac{2}{\pi} \arcsin\left(\frac{d}{\bar{d}}\right) - \frac{1}{\pi} \sin\left[2\arcsin\left(\frac{d}{\bar{d}}\right)\right]. \tag{29.40}$$

The above formula shows the effect of the off-axis light spot on the mixing efficiency, and the integral area is the overlapping area of the two light spots. In the formula, s is the area of the part where the signal light and the local oscillator light spot overlap, $\overline{d}$ is the radius of the overlapping area of the light spot, and d_1 and d_2 are the signal light and the light spot size of the local oscillator, respectively.

The mixing efficiency is attenuated by an increase in the angular deviation of the optical axis of the two light beams caused by the shaking of the visual axis. Compared with several other factors, the influence of the focal plane offset of the light spot caused by the shaking of the visual axis on the mixing efficiency is greatly improved. Good coaxiality of the optical axes of the two laser beams at the receiving end is very important. It can be compensated by improving the dynamic tracking of the optical signal and the aiming accuracy between the signal transmission and the receiving system.

Polarization Characteristics

The optical mixer can be regarded as a fixed-purpose polarization interferometer composed of a polarization-splitting prism and a waveplate. It mixes the signal with a constant phase difference, polarization matching, and close frequency with the local optical wavefront on the detector surface. The signal light gains by interfering with the local oscillator light, thereby improving the sensitivity of coherent

detection. Therefore, polarization matching is an important light beam coherence condition [178].

In a coherent optical communication system, before the laser beam is incident on the mixer, the polarization-maintaining fiber can ensure the stability of the polarization state of the local oscillator; however, the polarization state of the signal light is easily affected after passing through the complex spatial-transmission channel and optical transmission system. Therefore, the influence of the polarization characteristics of the signal light on the performance of the mixer is studied [179].

Influence of the Optical Vector Direction of the Linear Polarization Signal on the Mixing Efficiency

After the linearly polarized signal light and the local oscillator light are incident on the optical mixer, the DC branch photocurrent, used to restore the modulation information, and the quadrature branch photocurrent, used to control the loop, are output as follows:

$$I_{\mathrm{I}} = 2k_2k_3|E_{\mathrm{S}}E_{\mathrm{LO}}|\cos\phi(t) \tag{29.41}$$

$$I_{\mathrm{Q}} = 2k_1k_4|E_{\mathrm{S}}E_{\mathrm{LO}}|\cos\left(\phi(t) + 2\left(\tau_{//} - \rho_{\perp}\right) + \frac{\pi}{2}\right). \tag{29.42}$$

In the formula, k_1, k_2, k_3, and k_4 are the polarization vectors of the signal light and the polarized light, I_{I} is the modulation information reduction photocurrent, and I_{Q} is the photocurrent that controls the quadrature branch. Considering its impact on the performance of the mixer from the perspective of output power, the mixing-efficiency formula is written as [180]

$$\eta(\theta) = \frac{P_{\mathrm{Actual}}}{P_{\mathrm{Ideal}}} \propto \frac{I^2_{\mathrm{Actual}}}{I^2_{\mathrm{Ideal}}} = \frac{\left(\sqrt{2}\sin\theta + \sqrt{2}\cos\theta\right)^2}{4} = \frac{|1 + \sin 2\theta|}{2}. \tag{29.43}$$

Influence of Circular Polarization and Elliptical-Polarization Signal Lights on the Mixing Efficiency

Considering that circularly polarized and elliptically polarized light have better stability than the linearly polarized light used in traditional schemes, they have stronger adaptability to the atmosphere and have good polarization retention, especially in satellite-to-ground links. It is beneficial to improve the certainty of the polarization state of the communication light [181]. The situation when signal lights with different polarization states is incident can be simulated.

Therefore, in the case of ideal phase difference, as above, from the perspective of output power, considering its impact on the performance of the mixer, the mixing-efficiency formula is written in the form of an output power ratio:

$$\eta(\phi(t)-\psi,\theta)=\frac{I_{\text{Actual}}^2}{I_{\text{Ideal}}^2}=\frac{[\cos(\phi(t)-\psi)+\sin(\phi(t)-\psi)\sin(2\theta)]^2}{2} \tag{29.44}$$

$$\eta_c=\frac{\left[\int_0^{r_0}E_{\text{S}}(r)E_{\text{LO}}(r)\exp(i\Delta\varphi)rdr\right]^2}{\int_0^{r_0}[E_{\text{LO}}(r)]^2rdr\int_0^{\infty}[E_{\text{S}}(r)]^2rdr}\xrightarrow{\text{Energy changes and phase matches}}\frac{\left[\int_0^{r_0}\eta(\phi(t)-\psi,\theta)E_{\text{S}}(r)E_{\text{LO}}(r)rdr\right]^2}{\int_0^{r_0}[E_{\text{LO}}(r)]^2rdr\int_0^{\infty}[E_{\text{S}}(r)]^2rdr}. \tag{29.45}$$

When the angle between the fast axis of the quarter waveplate and the horizontal direction is 45°, the signal light entering the mixer is in a circular polarization state. At this time, the output signal power is the largest, which can be regarded as the relative mixing efficiency reaching its peak. When the fast axis of the quarter waveplate is located at other positions, the signal light is elliptically polarized, and the mixing efficiency decreases with the change of the angle θ of its fast axis.

Research on Optical Mixers at Xi'an University of Technology

Since 2000, Xi'an University of Technology has conducted in-depth theoretical research and experimental verification in complex environments and systems, and achieved good results in key areas of optical communication technologies, such as the following:

1. Transmission characteristics in laser atmospheric turbulence
2. Codecs in wireless laser communication systems
3. Optical wireless subcarrier modulation
4. Laser modulation and demodulation
5. Channel estimation and equalization
6. Design of optical transceiver systems
7. Acquisition-alignment tracking technology
8. Optical orthogonal frequency division multiplexing modulation and optical multiple input multiple output system research and
9. Optical wireless transmission theory and technology [182].

The Optoelectronic Technology Laboratory of Xi'an University of Technology conducts systematic research on many key elements of spatial coherent optical communications systems including the modulation mode, balanced detection module, adaptive-optics module, etc. A self-designed double-balanced detection receiver module is used in the self-designed and developed spatial coherent optical

communication system, and the performance of the detection system is theoretically analyzed and experimentally verified.

According to the crystal birefringence effect and comprehensively considering the function of the optical mixer, the Optoelectronic Technology Research Laboratory of Xi'an University of Technology designed a 2 × 4 90° crystal spatial optical mixer; analyzed its basic working principle, and produced its parameter design scheme. Finally, by establishing an error analysis model of the mixer, the phase error of the mixer, the uniformity of the branch energy, and the mathematical relationship between the I/Q splitting ratio of the mixer and the influencing factors were obtained.

Structural Design of an Optical mixer

Considering the function of the optical mixer and the optical properties of the uniaxial crystal, the schematic diagram of a crystal 2 × 4 90° spatial optical mixer is designed, as shown in Fig. 29.22. In Fig. 29.22, E_S and E_L represent the signal light beam and the local oscillator beam, respectively; all propagate along the z-axis. Along the light propagation direction are a quarter waveplate, birefringent crystal 1, first half waveplate, second half waveplate, birefringent crystal 2, and birefringent crystal 3.

In this structure, the main cross sections of crystals 1 and 2 are perpendicular to each other, the main cross sections of crystals 2 and 3 are at 45°, and the structures of crystals 1, 2, and 3 are the same. The optical axis x_3 is in the XOZ plane and forms an angle θ with the z-axis; x_2 is parallel to the y-axis. The surface of the quarter waveplate of the local oscillator optical path is parallel to the incident surface of crystal 1, and the fast axis of the waveplate is parallel to the x-axis. The surfaces of the first and second half waveplates are parallel to the incident surface of crystal 2, and the fast axis of the two waveplates is 45° with the x-axis.

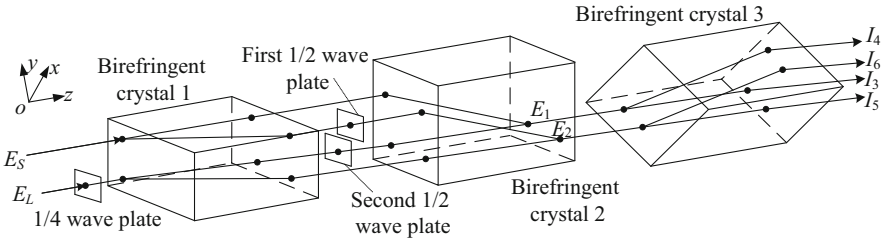


Fig. 29.22 Principle structure diagram of a crystal type 2 × 4 90° spatial light mixer (E_S: signal light beam; E_L: Local light beam; I_3, I_4, I_5, and I_6: four mixed output light beams)

Structural Optimization

The phase difference and splitting ratio are the most important factors to measure the performance of an optical mixer. Ideally, the phase difference of the I/Q of the mixer is 90° and the splitting ratio is 1. However, in practice, owing to the influence of various factors, errors will occur in the phase and splitting ratio of the mixer, which will affect its performance.

During the principle analysis of the mixer, it was found that the phase shift required by the mixer was provided by the quarter waveplate. Rotating the optical axis of the quarter waveplate can compensate the phase of the mixer; however, the resulting side effect is that the splitting ratio between the I/Q channels is changed.

As long as the Q branch can be properly locked, more energy can be distributed to the I branch for demodulating the signal, thereby improving the coherent detection sensitivity; hence, a method is needed to adjust the splitting ratio. A half waveplate is added to the signal optical path of the prism-type spatial optical mixer, and the splitting ratio can be adjusted by rotating this waveplate. It was applied to the optical mixer designed in this study.

The experimental results of the adjusted optical mixer show that when the laser frequency difference is greater than 400 MHz, it has a severe impact on the phase accuracy of the mixer; the quarter waveplate has a severe impact on the phase and splitting ratio of the optical mixer; and the crystal transmittance has a severe impact on the branch energy uniformity and splitting ratio. The phase compensation and splitting ratio adjustment scheme is simple and effective and has little impact on the mixing efficiency of the mixer.

Experimental Verification

The experimental system with the structure shown in Fig. 29.23 was built, and the designed crystal optical mixer was tested.

Through testing, it was found that two sine waves appeared on the oscilloscope. As shown in Figs. 29.24 and 29.25, the quarter waveplate is rotated for debugging, so that the amplitude between the two electrical signals is about 1:1. This verifies that

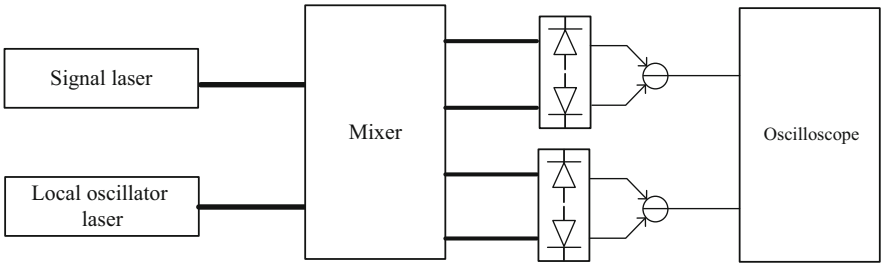


Fig. 29.23 Test-system structure diagram

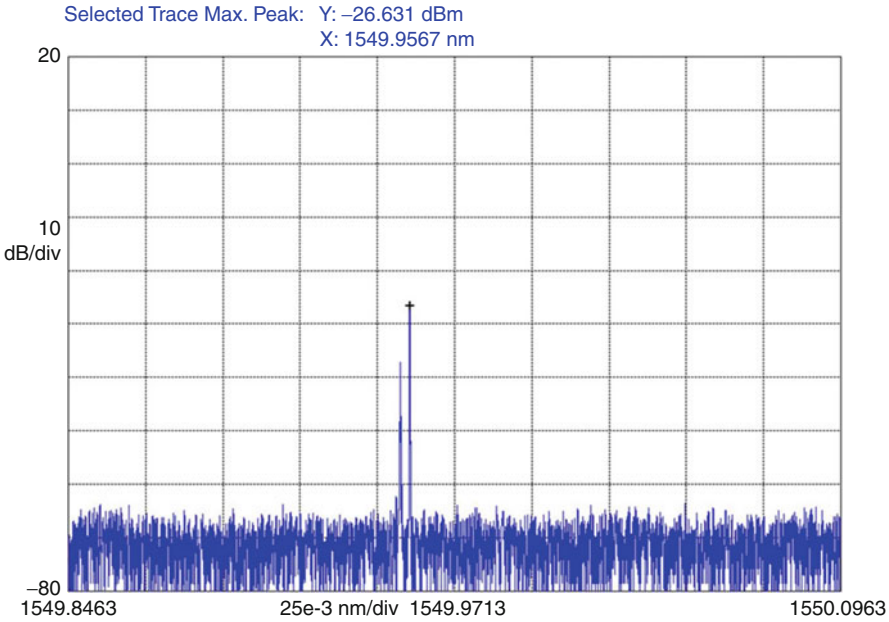


Fig. 29.24 Spectrum of the output intermediate frequency signal

the mixer can mix the signal light and local oscillator light. At this time, the intermediate frequency voltage of the two signals reaches between 130 mV and 145 mV. Because the voltage value detected by the detector is proportional to the received optical power, the insertion loss of the designed optical mixer is small.

Finally, the designed optical mixer is applied to the BPSK coherent optical communication system. The camera loads the collected signal into the optical signal through the phase modulator and transmits it to the optical mixer through the single-mode optical fiber. After the received signal light is mixed with the local oscillator light, the obtained quadrature branch controls the frequency of the local oscillator light through the frequency control system.

After a photoelectric conversion and differential processing are carried out by the balanced detector, the digital signal processing unit demodulates the signal, and the source is displayed on the computer as a video. After testing, the mixer can be used in coherent optical communication systems normally and can carry out long-term stable communication, which further verifies the function of the mixer.

Summary

This chapter systematically introduced the research status and progress of optical mixers and research results and compared the characteristics, advantages, and disadvantages of different structures and types. A high signal-to-noise ratio and



Fig. 29.25 Crystal optical mixer intermediate frequency signal test results

miniaturized, integrated, and high-stability optical mixers are still the goals pursued by researchers. They are also the future development trends of optoelectronic detector devices.

1. Coherent optical communication must continuously meet the rapid and normal communication needs of an information society, and continuously expand its own communication capacity and bandwidth. As a key component of a coherent optical communication system, the performance of the optical mixer is a key issue that must be solved.

   At present, coherent optical communication is concentrated in the C band and L band; thus, finding more communication bands has become a research hotspot for the next generation. The corresponding optical mixer should also be compatible with more bands, expand its bandwidth, and improve spectrum utilization, which are inevitable trends in the development of coherent optical communication.
2. Traditional optical mixers are mostly block and optical element structures, which are inconvenient for integration development. Therefore, the integrated development of optical mixers is an inevitable trend. Although the development of materials has promoted the integrated development of optoelectronic devices, there are still many challenges. The high capacity of new wave bands and the applicability of various high-order modulation formats should be adapted. Therefore, the integrated development of optical mixers is of great significance for developing new wave bands and realizing compact coherent light receivers.

   Existing optical modules are mainly based on silicon optical coupling technology, including optical mixers, and various transmit and receive optical devices. However, many problems in its development limit the communication capacity. The large waveguide loss and high insertion loss are problems to be solved. Therefore, improving the performance of silicon-based optical mixers is an effective way to further improve the communication capacity and solve current problems.

In the future, with the increasing integration of photons and the development of basic disciplines, such as material science and electronic science, both optical mixers and balanced detectors will eventually be integrated, through integrated technologies and more complex manufacturing processes, to establish high-performance photoelectric detection platforms.

Mixer research mainly focuses on two aspects. It improves the performance and reduces the signal-to-noise ratio by continuously designing new mixer structures, and it investigates the theoretical relationship between the various mixer parameters. With the continuous improvement of theoretical research work and the continuous emergence of new structures, mixer performance has made a huge leap and has been applied in many fields. As a key device in coherent optical communication, the mixer has received extensive attention and research. In the future, the exploration of new structures and improving the common mode rejection ratio and signal-to-noise ratio are still the goals of researchers.

The continuous optimization and improvement of photodetector devices will further promote the application and development of the coherent optical communication, photoelectric detection, and lidar fields. Coherent optical communication technology will further improve its communication capacity and communication quality, owing to the development of coherent optoelectronic devices, and better provide high-quality communication services for information societies.

References

1. Armstrong EH. The superheterodyne-its origin, development, and some recent improvement [J]. Pro. IRE, 1924, 3(12):540–552
2. Xie Jinjuan. Design and implementation of high and medium frequency circuit at the front end of shortwave receiver [D]. Wuhan: Wuhan University of Technology, 2010.
3. Li Chunguang, Wang Jia, Wu Yun, et al. Research and application progress of superconducting electronics in China [J]. Chinese Journal of Physics, 2021, 70 (1): 184–209.
4. Wschottky. Uber super heterodyngedankens [J]. Elek. Nachrichtentech, 1925, 5(2): 48–54
5. Song Muxiang. The inventor of the superheterodyne circuit - Edwin H. Armstrong [J]. Chinese Engineer, 1995, (04): 33–35.
6. Hu Haifan, Zhao Qingran, Ma Xuming, et al. Research on 250 GHz second harmonic mixer of GaAs Schottky diodes [J]. Infrared and Laser Engineering, 2019, 48 (7): 12–18.
7. Wen Z, Deng L, Zhao X, et al. Improving voltage output with PZT beam array for MEMS-based vibration energy harvester: Theory and experiment [J]. Micro system Technologies, 2015, 21(2):331–339.
8. Wei Ping. Research on ultra-wideband microwave mixer [D]. Chengdu: University of Electronic Science and Technology of China, 2008.
9. He Tingting, Li Shaofu, Mahmud Nahid-Al, et al. Research progress of Schottky diode terahertz mixers [J]. Sensors and Microsystems, 2020, 39 (5): 4–6.
10. He Jie, 1THz High-order Harmonic Mixer [D]. Chengdu: University of Electronic Science and Technology of China, 2018.
11. He Rong. Mixer design for millimeter-wave radar receiver [D]. Beijing: Beijing Information Science and Technology University, 2021.
12. Feng Zhenyu. Research on terahertz solid-state mixer based on planar Schottky barrier diode [D]. Chengdu: University of Electronic Science and Technology of China, 2019.
13. Thomas B, Maestrini A, Beaudin G. A low-noise fixed-tuned 300–360 GHz sub-harmonic mixer using planar Schottky diodes [J]. Microwave Wireless Components Letters, 2005, 15(12):865–867.
14. J. Zhang. Wide-band equivalent circuit of Quasi-Vertical Planar Schottky Diode for 650 GHz Subhammonic mixers. st [J]. IRAM International Symposium On THz Electronics, 1997, 35(19):223–226
15. J. L. Hesler. Planar Schottky diodes in submillimeter-wavelength waveguide receivers [J]. Space THz Technology, 1993, 13(2):12–21.
16. J. L. Hesler. A Fixed tuned 400GHz Subharmonic Mixer Using Planar Schottky Diodes [J]. Space THz Technology, 1999, 78(1):95–99.
17. PH Siegel, Smith, R. P, Graidis, M. C. et al. 2.5-THz GaAs monolithic membrane-diode mixer IEEE transactions [J]. Microwave Theory and Techniques, 1999, 47(5):596–604.
18. Giustolisi G, Palmisano G, Palumbo G, et al. A novel 1.5-V cmos mixer [C]. Proceedings of the 8th Great Lakes Symposium on VLSI (Cat. No. 98TB100222). IEEE, Lafayette, LA, USA, 1998: 113–119.
19. U. V. Bhapkar, Y. Li, R. J. Mattauch. InGaAs/InP heteroepitaxial Schottky barrier diodes for Terahertz ap plications [J]. International Symposium on Space THz Technology, 1992, 32(1): 661–677.

20. F. Podevin. Composants Schottky a heterostructures de semiconducteurs en technologie InP pourle melange de frequences a 560 GHz [D]. Lille: Universite des Sciences et Technologier, 2001.
21. J. L. Hesler, D. W. Porterfield, W. L. Bishop, et al. Development of compact broadband receivers at submillimeter wavelengths [J]. Aerospace Conference, 2004, 22(1):735–740.
22. Fleri D A, Cohen LD. Nonlinear Analysis OFT the Schottky-Barrier Mixer Diode [J]. Transactions on Microwave Theory and Techniques, 1973, 21(1):39–43
23. Araki T, Hirayama M.A 20-GHz Integrated Balanced Mixer [J]. Transactions on Microwave Theory and Techniques, 1971, 19(7):638–643.
24. Mccoll M. Conversion Loss Limitation on Schottky-Barrier Mixers [J]. Transactions on Microwave Theory and Techniques, 1977, 25(1):54–59.
25. Wang Li. Design of K-band integrated down-converter [D]. Nanjing: Nanjing University of Science and Technology, 2008
26. Champlin K S, Eisenstein G. Cutoff frequency of submillimeter Schottky-barrier diodes [J]. IEEE Transactions on Microwave Theory and Techniques, 1978, 26(1): 31–34.
27. Adams J, Tang T W. A revised boundary condition for the numerical analysis of Schottky barrier diodes [J]. IEEE Electron Device Letters, 1986, 7(9): 525–531.
28. Jelenski A, Grub A, Krozer V, et al. New approach to the design and the fabrication of THz Schottky barrier diodes [J]. Transactions on Microwave Theory and Techniques, 1993, 41(4): 549–557.
29. Lipsey R E, Jones S H. Monte Carlo harmonic-balance and drift-diffusion harmonic-balance analyses of 100–600 GHz Schottky barrier varactor frequency multipliers [J]. Transactions on Electron Devices, 1997, 44(11): 1843–1850.
30. Hu Jiankai, Huang Zhenqi, Li Baoling, et al. Research on fin line mixers in 3mm FM/CW modules [J]. Research and Progress in Solid State Electronics, 2001, 21 (1): 4–5.
31. Hong Qi. Design of CMOS low-voltage high-performance mixer [D]. Hefei, Anhui: Anhui University, 2005.
32. Yang Dejun. Design of CMOS RF Linear Mixer [D]. Chengdu: University of Electronic Science and Technology of China, 2006.
33. Wu Rui. Noise analysis of RF integrated active mixer [D]. Chengdu: University of Electronic Science and Technology of China, 2009.
34. Cooper K B, Dengler R J, Llombart N, et al. THz Imaging Radar for Standoff Personnel Screening [J]. IEEE Transactions on Terahertz Science and Technology, 2011, 1(1):169–182.
35. Siles J V, Grajal J, Carlo A D. Design of submillimeter Schottky mixers under flat-band conditions using an improved drift-diffusion model [J]. Microwave and Wireless Components Letters, 2009, 19(3): 167–169.
36. Tian Chao, Yang Hao, Dong Junrong, et al. Design and fabrication of an exponentially doped GaAs planar Schottky varactor diode [J]. Electronic Devices, 2011, 21 (1): 29–32.
37. Saber M A. Thermal characterization of THz planar Schottky diodes using simulations [D]. Finland: Aalto university, 2012.
38. Yan L, Krozer V, Michaelsen R, et al. Physical based Schottky barrier diode modeling for THz applications [C]. IEEE International Wireless Symposium, Beijing, China, 2013: 1–4.
39. Pérez-Moreno C G, Grajal J. Physical electro-thermal model for the design of Schottky diode-based circuits [J]. Transactions on Terahertz Science and Technology, 2014, 4(5):597–604.
40. Orfao B, Vasallo B G, Moro-Melgar D, et al. Analysis of surface charge effects and edge fringing capacitance in planar GaAs and GaN Schottky barrier diodes [J]. Transactions on Electron Devices, 2020, 67(9): 3530–3535.
41. Orfao B, Vasallo B G, Pérez S, et al. Dielectric passivation and edge effects in planar GaN Schottky barrier diodes [J]. Transactions on Electron Devices, 2021, 68(9): 4296–4301.
42. Tang Deyao, Zhou Manling. CMOS analog switching mixer [J]. Electric Measurement and Instrument, 1984, 22 (06): 35–36.
43. Toole W H, El-Masry E I, Manku T. A novel highly linear 1 GHz switched-current sub--sampling mixer [C]. 1996 IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World. ISCAS 96. IEEE, Atlanta, GA, USA, 1996, 1: 203–206.

44. Wang H M. A 1 V multi-gigahertz RF mixer core in 0.5/spl mu/m CMOS [C]. 1998 IEEE International Solid-State Circuits Conference. Digest of Technical Papers, ISSCC. First Edition (Cat. No. 98CH36156). IEEE, San Francisco, CA, USA, 1998: 370–371.
45. Li L, Guo J, Li Z, et al. Substrate noise analysis for RF CMOS mixers based on state equation technique [C]. Proceedings of the Third International Workshop on Design of Mixed-Mode Integrated Circuits and Applications (Cat. No. 99EX303). IEEE, Puerto Vallarta, Mexico, 1999: 94–97.
46. Terrovitis M T, Meyer R G. Noise in current-commutating CMOS mixers [J]. Journal of solid-state circuits, 1999, 34(6): 772–783.
47. Chi Baoyong, Shi Bingxue. CMOS mixer for 2.4GHz wireless local area network transceiver [J]. Chinese Journal of Semiconductors, 2003, 12 (05): 472–475.
48. Terrovitis M T, Meyer R G. Intermodulation distortion in current-commutating CMOS mixers [J]. Journal of Solid-State Circuits, 2000, 35(10): 1461–1473.
49. El-Shewekh E W, El-Saba M H. Wide frequency range (30 M-1 GHz) low noise figure, low power, active CMOS mixer [C]. Proceedings of the 12th IEEE International Conference on Fuzzy Systems (Cat. No. 03CH37442). IEEE, Cairo, Egypt, 2003: 160–163.
50. Cui Fuliang. Analysis and Design of Mixer Circuit and Image Suppression Circuit in Bluetooth Transceiver [D]. Shanghai: Fudan University, 2004.
51. Ren Lening, Zhu Zhangming, Yang Yintang. A 0.6V, 2.4 GHz CMOS mixer based on quasi-floating gate technology [J]. Journal of Circuits and Systems, 2005, 6 (01): 144–146.
52. Dong Huaiyu. Research on CMOS RF UWB Mixer [D]. Xi'an: Xi'an University of Technology, 2005.
53. Liu Qiankun. Design and research of CMOS direct down-conversion mixer in ultra-wideband system [D]. Xi'an: Xi'an University, 2007.
54. Zhou Zhizeng, Li Dongsheng. Design of a 0.18 μm CMOS wideband image suppression mixer [J]. Information and Electronic Engineering, 2008 (01): 17–23.
55. Wei Changming, Chen Diping, Chen Qixing. Design of a 900MHz CMOS low-voltage high linearity mixer [J]. Semiconductor Technology, 2005, 56 (10): 78–82.
56. Wu Mingming. Design and Simulation of 2.4GHz CMOS Mixer [D]. Harbin: Harbin Institute of Technology, 2006.
57. Tang Shoulong, Wu Jianhui, Luo Lan. Noise analysis of CMOS Gilbert mixer in deep submicron process [J]. Chinese Journal of Applied Science, 2006, 6 (04): 372–376.
58. Yu Haoran. Analysis and Design of 0.35 μ m Bi CMOS Mixer in 2.4GHz Receiver Chip [D]. Chengdu: University of Electronic Science and Technology of China, 2007.
59. Yuan Shuaijun. Research on CMOS current-mode radio frequency mixer [D]. Changsha: Hunan University, 2010.
60. Zhang Bang, Song Zhiguo, Deng Xiaofei. Design of low-voltage and high-gain down-conversion mixer [J]. Microcomputer Information, 2012, 28 (09): 183–184
61. Shen Jing, Zhang Xiaolin. Ultra-low power consumption current multiplexing mixer based on substrate bias [J]. Beijing: Journal of Beihang University, 2013, 39 (04): 531–534.
62. Hu Junwei. Research and design of millimeter wave mixer based on CMOS process [D]. Chengdu: University of Electronic Science and Technology of China, 2013.
63. Wang Chao. Design of millimeter-wave CMOS Schottky barrier diode mixer [D]. Tianjin: Tianjin University, 2014.
64. Wegerif S, Redman-White W. An integrated CMOS image-rejection mixer system for low-jitter secondary frequency references [C]. 1993 IEEE International Symposium on Circuits and Systems. IEEE, Chicago, IL, USA, 1993: 1010–1013.
65. Hooman Darabi and Asad A. Abide "Noise in RF-CMOS Mixers: A Simple Physical Model" [J]. Transactions on Solid-State Circuits, 2013, 39(04): 531–534.
66. Yang Xiaofan, Research on terahertz sub-harmonic mixer based on planar Schottky diode [D]. Chengdu: University of Electronic Science and Technology of China, 2012.

67. Bulcha B T, Hesler J L, Drakinskiy V, et al. Design and characterization of 1.8–3.2 THz Schottky-based harmonic mixers [J].Transactions on Terahertz Science and Technology, 2016, 6(5):737–746.
68. Nedungadi A, Viswanathan T. Design of linear CMOS transconductance elements [J]. Circuits & Systems IEEE Transactions on, 1984, 31(10):891–894.
69. Blaisdell A, Geo ffroy R, Howe H.A Novel Broadband Double Balanced Mixer for the 18-40GHz Range [J]. MTT-S Digest, 1982, 3(4):33–35
70. Ton T N, Dow G S, Chen T H, Lacon M, Lin T S, Bui S, Yang D. An X-Band Monolithic Double Double-Balanced Mixer for High Dynamic Receiver Application [J]. MTT-S Digest, 1990, 6(1):197–200
71. Eisenberg J, Panelli J, Weiming O.A New Planar Double-Double Balanced MMIC Mixer Structure [J]. MTT-S Digest, 1991, 2(1):81–84.
72. Salem R F, Galal S H, Tawfik M S. a new highly linear CMOS mixer suitable for deep submicron technologies. IEEE International Conference Electronics Circuit and Systems, 2002, (1):81–84
73. Hung-Che Wei, Ro-Min Weng, Chih-Lung Hsiao, Kun-Yi Lin. A 1.5V 2.4 GHz CMOS mixer with high linearity [J]. Asia-Pacific Conference on Circuits and Systems, 2004(11):289–292
74. Shen Yihao, Zhang Runxi, Zhang Weijie, et al. Analysis and application of multi-hyperbolic tangent rule in the design of high linearity CMOS mixers [J]. Electronic Devices. 2008, 31 (4): 1150–1154
75. V. Aparin, G. Brown, L. E. Larson. Linearization of CMOS LNAs via optimum gate biasing [J]. Int. Circuits Syst. Symp. 2004, 1(04):748–750.
76. D. R. Webster et al, Derivative superposition -a linearization technique for ultra-broadband systems [J]. Colloq. On Wideband Circ. Modeling Tech. 1996, 3(1):12–14
77. Tae Wook Kim, Bonkee Kim, Kwyro Lee. Highly linear receiver Front-end adopting MOSFET transconductance linearization by multiple gated transistors [J]. Solid-State Circuits. 2004, 39(1):223–229.
78. Aparin V, Larson L E. Modified derivative superposition method for linearizing FET low-noise amplifiers [J]. IEEE Transactions on Microwave Theory & Techniques, 2005, 53(2):571–581.
79. Cusmai G Brandolini M, Rossi P.A 0.18-um CMOS Selective Receiver Front-End for UWB Application [J]. Journal of Solid-State Circuits. 2006, 41(1):1764–1771
80. Benisty M, Berger J, Jocou L, et al. An integrated optics beam combiner for the second generation VLTI instruments [J]. Astronomy and Astrophysics, 2009, 498(2):601–613.
81. Deng C, Gong W, Han S. Pulse-compression ghost imaging lidar via coherent detection [J]. Optics Express, 2016, 24(23):25983–25994.
82. Yan D, Popp J, Pletz M, et al. Highly sensitive broadband Raman sensing of antibiotics in step-index hollow-core photonic crystal fibers [J]. ACS Photonics, 2017, 4(1):139–145.
83. Chen Hao. Design and optimization of silicon dioxide-based 90-degree optical mixer [D]. Hangzhou, Zhejiang: Zhejiang University, 2021.
84. Ma Tingting. Analysis of the influence of external factors on the mixing efficiency of spatial optical mixers [D]. Changchun: Changchun University of Science and Technology, 2019.
85. Kazovsky L G, Curtis L, Young W C, et al. All-fiber 90 optical hybrid for coherent communications [J]. Applied optics, 1987, 26(3): 437–439.
86. Berenbrock G A, Schlemmer B. Active controlled fiber optical 90 degrees hybrid for coherent communications [J]. Photonics Technology Letters, 1989, 1(4): 86–87.
87. Stowe D, Hsu T Y. Demodulation of interferometric sensors using a fiber-optic passive quadrature demodulator [J]. Journal of Lightwave Technology, 1983, 1(3): 519–523
88. Yuen H P, Chan V W. Noise in homodyne and heterodyne detection [J]. Optics Letters, 1983, 8(3): 177–179.
89. Abbas G L, Chan V W S, Yee T K. A dual-detector optical heterodyne receiver for local oscillator noise suppression [J]. Journal of Lightwave Technology, 1985, 3(5): 1110–1126.

90. Gimlett J, Vodhanel R, Choy M, et al. A 2-Gbit/s optical FSK heterodyne transmission experiment using a 1520-nm DFB laser transmitter [J]. Journal of lightwave technology, 1987, 5(9): 1315–1328.
91. Inoue T, Nara K. Ultrasmall PBS-integrated coherent mixer using 1.8%-delta silica-based planar light-wave circuit [J]. Optical Communication, 2010, 23(1): 23–30.
92. Luo Yuchan. Research on silicon-based integrated taper directional coupler suitable for multi-dimensional multiplexing systems [D]. Wuhan: Huazhong University of Science and Technology, 2019.
93. Chen B. X, Lu H. L, Zhao D. X, et al. Optimized design of polarization-independent and temperature-insensitive broadband optical waveguide coupler by use of fluorinated polyimide [J]. Applied Optics, 2003, 42(20): 4196–4201.
94. Wang Q., He S. Optimal design of planar wavelength circuits based on Mach-Zehnder Interferometers and their cascaded forms [J]. IEEE Journal of Lightwave Technology, 2005, 23(3): 1284–1290.
95. Alam M. Z., Caspers J. N., Aitchison J. S., et al. Compact low loss and broadband hybrid plasmonic directional coupler [J]. Optics Express, 2013, 21(13): 16029–16034.
96. Zhou Y, Sun J, Yan A, et al. Self-homodyne interferometric detection in 2× 4 optical hybrid for free-space optical communication [J]. Laser Communications, 2010, 26(2): 65–78.
97. Tang Z, Zhang F, Zhu D X, et al. A photonic frequency downconverter based on a single dual-drive Mach-Zehnder modulator [J]. IEEE Topi-Cal Meeting Microw. Photon., 2013, 42(12): 150–153.
98. Zhang Xiaoyu, Qiu Xinjing, He Ruiying, et al. All-optical mixer based on dual parallel Mach-Zehnder modulators [J]. Optical Communication Technology, 2019, 43 (1): 47–50.
99. Garreis R B. 90-degree optical hybrid for coherent receivers [J]. SPIE, 1991, 15(22): 210–221.
100. Ma Tingting, Tong Shoufeng, China Southern Airlines, et al. Influence of signal optical polarization characteristics on mixing efficiency of spatial coherence detection [J]. Advances in Laser and Optoelectronics, 2017, 54 (2): 110–118.
101. Barwicz T., Watts M. R., Popovi, et al. Polarization-transparent microphotonic devices in the strong confinement limit [J]. Nature Photonics, 2006, 1(1): 57–60.
102. Dai D, Liu L, Gao S, et al. Polarization management for silicon photonic integrated circuits [J]. Laser & Photonics Reviews, 2013, 7(3): 303–333.
103. Teich M C. Infrared heterodyne detection [J]. Proceedings of the IEEE, 1968, 56(1): 37–46.
104. Kleinman D A, Boyd G D. Infrared detection by optical mixing [J]. Journal of Applied Physics, 1969, 40(2): 546–566.
105. Langenhorst W. Compact bulk optical 90 degrees hybrid for balanced phase diversity receivers [J]. Electronics Letters, 1989, 25(22):1518–1521.
106. Garreis R B, Zeiss C. 90° optical hybrid for coherent receivers [J]. Optics and Photonics, 1991, 15(22):210–219
107. Zhang Chenxiang, Chen Chao, Liu Zichen, et al. A 90° mixer for spatial optical input [J]. Optical Communication Research, 2016, 193 (01): 45–47.
108. Inoue T, Kawashima H, Nara K. Double-pass PBS-integrated coherent mixer using silica-based PLC [J]. Electronics Letters, 2010, 26(21):1517–1521.
109. Zhai Yumeng, Wang Jin, Xu Ji, et al. Large tolerance coherent optical mixer based on polymer multimode interferometer [J]. Optoelectronics Laser, 2015, 26 (2): 233–238
110. Zhuang Xinxin, Su Genbo, He Youping, et al. Birefringent crystal - an important optical functional material [J]. Structural Chemistry, 2001, 21 (4): 291–293.
111. Li Gangsheng, Wu Xiquan. Growth of large-scale high-quality yttrium vanadate (YVO4) birefringent crystal [J]. Journal of Artificial Crystals, 1999, 28 (1): 27–30.
112. Wang Yingwei, Cheng Haobo, Liu Jinghe. Growth and integrity analysis of YVO4 birefringent crystal [J]. Journal of Artificial Crystals, 2013, 32 (1): 41–43.
113. Wu H, Xu H, Su Q, et al. Size and shape tailored hydrothermal synthesis of YVO4 crystals in ultra-wide pH range conditions [J]. Journal of Materials Chemistry, 2003, 13(5):1223–1233.

114. Rybaromanowski W. YVO4 crystals puzzles and challenges [J]. Crystal Research & Technology, 2010, 38(3–5):225–236.
115. Li Gangsheng, Wei Min, Zhu Yuemei, et al. Dehydration synthesis of yttrium vanadate birefringent crystal raw materials [J]. Chinese Journal of Artificial Crystals, 1997, 26 (3–4): 400–405.
116. Li J, Billah M R, Schindler P C, et al. Four-in-one interferometer for coherent and self-coherent detection [J]. Optics express, 2013, 21(11): 13293–13304.
117. Wan Lingyu, Sun Jianfeng, Liu Liren, Influence of wave plate optical axis deviation on amplitude and phase of light beam with polarization splitting [J], Optical Technology, 2010, 21 (36): 307–310
118. Ke Changjian, Qian Lu, Wan Zhujun. A coherent receiver optical mixer [P]. Chinese patent: CN103166713A.2013.
119. Zheng Yang, Jiang Huilin, Tong Shoufeng, et al. Establishment of mathematical model of spatial optical mixer based on coherent laser communication [J]. Chinese Journal of Optics, 2013, 33 (7): 149–154.
120. Song Zhe, Yu Weixing, Song Jia, et al. 2 × 4 integrated structure 90° space optical bridge [J]. Advances in Laser and Optoelectronics, 2018, 55 (11): 438–443.
121. Han Jianlu. Design and research of crystal type 2 × 4 90° spatial optical mixer [D]. Xi'an: Xi'an University of Technology, 2023.
122. Zhang Lin. Research on the integration of space optical bridges [D]. Shenyang: Liaoning Normal University, 2019.
123. Nicholson G, Stephens T D. Performance analysis of coherent optical phase-diversity receivers with DPSK modulation [J]. Journal of Lightwave technology, 1989, 7(2): 393–399.
124. P. Dong, C. Xie, L.L. Buhl. Monolithic polarization diversity coherent receiver based on 120-degree optical hybrids on silicon [J]. Optics Express, 2014, 22(2): 2119–2125
125. Doerr C R, Gill D M, Gnauck A H, et al. Monolithic demodulator for 40·Gb/s DQPSK using a star coupler [J]. Journal of Lightwave Technology, 2006, 24(1):171–174.
126. Doerr C R, Zhang L, Winzer P J. Monolithic InP multiwavelength coherent receiver using a chirped arrayed waveguide grating [J]. Journal of Lightwave Technology, 2011, 29(4): 536–541.
127. Zimmermann L, Voigt K, Winzer G, et al. C—Band optical 90° hybrids based on silicon-on-insulator 4×4 waveguide couplers [J]. IEEE Photonics Technology Letters, 2009, 21(3): 143—145.
128. Jeong S H, Morito K. Compact optical 90° hybrid employing a tapered 2x4 MMI coupler serially connected by a 2x2 MMI coupler [J]. Optics Express, 2010, 18(5):4275–4288
129. Sakamaki Y, Nasu Y, Hashimoto T, et al. Reduction of phase-difference deviation in optical hybrid over wide wavelength range [J]. IEICE Electronics Express, 2010, 7(3):216–221.
130. Guan H, Ma Y, Shi R, et al. Compact and low loss 90° optical hybrid on a silicon-on-insulator platform [J]. Optics Express, 2017, 25(23):28957–28968
131. Yang W, Yin M, Li Y P, et al. Ultra-compact optical 90° hybrid based on a wedge-shaped 2×4MMI coupler and a 2×2MMI coupler in silicon-on—insulator [J]. Optics Express, 2013, 21(23):28423–2843.
132. Yin M, Yang W, Huang Y, et al. Compact and wideband optical 900 hybrid based on silicon—on—insulator [J]. International Photonics. 2013, 1(02):57–58.
133. Voigt K, Zimmermann L, Winzer G, et al. C-Band optical 90° hybrids in silicon Nano waveguide technology [J]. IEEE Photonics Technology Letters, 2011, 23(23):1769–1771.
134. Jeong S, Morito K. Compact and wideband optical 90° hybrid based on a one-way tapered MMI coupler [C]. 2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, 2011:1–3.
135. Xu L, Wang Y, Patel D, et al. Ultra—broadband and ultra—compact optical 90° hybrid based on 2×4MMI coupler with subwavelength gratings on silicon-on-insulator [C]. 2018 Optical Fiber Communications Conference and Exposition(OFC). 2018:1–3.

136. Wang Z, Zhai Y, Lu Y, et al. Compact optical 90° hybrid based on hybrid plasmonic multimode interferometer [J]. Optics Communications, 2018, 42(6):99–104.
137. Jiao Y, Zhu Y, Hong X, et al. An integrated optical mixer based on SU8 polymer for PDM—QPSK demodulation [J]. IEEE Photonics Technology Letters, 2011, 23(20):1490–1492
138. Yu J, Mu J, Chen K, et al. High-performance 90。hybrids based on MMI couplers in Si3N4 technology [J]. Optics Communications, 2020, 46(5):125–129.
139. Nasu Y, Mizuno T, Kasahara R, et al. Temperature insensitive and ultrawide band silica-based dual polarization optical hybrid for coherent receiver with highly symmetrical interferometer design [J]. Optics Express, 2011, 1 9(26):BI12–8.
140. Jeong S-H, Morito K. Optical 45° hybrid for demodulating 8-ary DPSK signal [J]. Optics Express, 2010, 18(8):8482–8490.
141. Jeong S-H, Morito K. Optical 60° hybrid for demodulating six-level DPSK signal [J]. Optics Letters, 2011, 36(3):322–324.
142. Saber M G, Osman M, Patel D, et al. Demonstration of a 120° hybrid based simplified coherent receiver on SOI for high speed PON applications [J]. Optics Express, 2018, 26(24): 31222–31232.
143. Li Z, Kuang Y, Guan H, et al. Compact low-loss optical 72° hybrid based on nonoverlapping. image multimode interference coupler in silicon-on-insulator [J]. IEEE Photonics Journal, 2019, 11(6):1–9.
144. Wang Jiayao. Research on silicon-based optical mixers [D]. Hefei: Chinese Academy of Sciences, 2024.
145. Sah P, Das B K. Photonic bandpass filter characteristics of multimode SOI waveguides integrated with submicron gratings [J]. Applied Optics, 2018, 57(9): 2277–2281.
146. Park J, Kwack M J, Joo J, et al. Low-loss single-mode operation in silicon multi-mode arrayed waveguide grating with a double-etched inverse taper structure [J]. Journal of Optics, 2017, 19(8): 1–10.
147. Hiraki T, Aihara, Nishi H, et al. Deuterated SiN/SiON Waveguides on Si Platform and Their Application to C-Band WDM Filters [J]. IEEE Photonics Journal, 2017, 9(5):1–7.
148. Yang Zhifang. Design and performance research of double-layer array waveguide grating based on polymer waveguide platform [D]. Harbin: Heilongjiang University, 2026.
149. Smit M K. New Focusing and Dispersive Planar Component Based on an Optical Phased Array [J]. Electronics Letters, 1988, 24(7): 385–386.
150. Vellekoop A R, Smit M K. A small-size polarization splitter based on a planar optical phased array [J]. Journal of Lightwave Technology, 1990, 8(1): 118–128.
151. Zirngibl M, Dragone C, Joyner C. Demonstration of a 15x15 arrayed waveguide multiplexes on InP [J]. IEEE Photonics Technology Letters, 1992, 10(1): 1250–1253.
152. Tong W, Menon V M, Xia F, et al. An asymmetric twin-waveguide eight channel polarization independent arrayed waveguide grating with an integrated photodiode array [J]. IEEE Photonics Technology Letters, 2004, 16(1): 1170–1172.
153. Hill M T, Vries T D, Dorren H, et al. Integrated two-state AWG-based multiwavelength laser [J]. Photonics Technology Letters, IEEE, 2005, 17(5): 956–958.
154. Heck M J R, La Porta A, Leijtens X J M, et al. Monolithic AWG-based Discretely Tunable Laser Diode with Nanosecond Switching Speed [J]. IEEE Photonics Technology Letters, 2009, 21(13): 905–907.
155. Ishii M, Takagi A, Hida Y, et al. Low-loss Fiber-pigtailed 256 channel arrayed-waveguide grating multiplexer using cascaded laterally-tapered waveguides [J]. Electronics Letters, 2001, 37(23): 1401–1402.
156. Stanton E J, Volet N, Bowers J. Low-loss arrayed waveguide grating at 2.0μm [J]. Optics Letters, 2017, 7(2): 1–2.
157. Zhang Z, Hu J, Chen H, et al. Low-crosstalk silicon photonics arrayed waveguide grating [J]. Chinese Optics Letters, 2017, 15(4): 1–4.

158. Song G. Y., Zou J., and He J. J. Ultra-compact silicon-arrayed waveguide grating routers for optical interconnect systems [J]. Chinese Optics Letters, 2017, 15(3): 1–4.
159. Stanton E J, Nicolas V, Bowers J E. Silicon arrayed waveguide gratings at 2.0-μm wavelength characterized with an on-chip resonator [J]. Optics Letters, 2018, 43(5): 1135–1138.
160. Elke, Pichler, Konrad, et al. Rapid prototyping of all-polymer AWGs for FBG readout using direct laser lithography [J]. Optics Letters, 2018, 43(21): 5347–5350.
161. Park J, Joo J, Kim G, et al. Low-Crosstalk Silicon Nitride Arrayed Waveguide Grating for the 800–nm Band [J]. IEEE Photonics Technology Letters, 2019, 31(14):
162. Han Q, Menard M, Shi W. Superlattice Arrayed Waveguide Grating in Silicon Nitride [J]. IEEE Photonics Technology Letters, 2020, 32(22): 1411–1414.
163. Liu J, Huang, Q. Z., Xia J. S., High Assembly Tolerance and Cost-Effective 100-Gb/s TOSA With Silica-PLC AWG Multiplexer [J]. IEEE Photonics Journal, 2019, 11(4): 1–9.
164. Zhang Z, Wang Y, Tsang H K. Ultracompact 40-Channel Arrayed Waveguide Grating on Silicon Nitride Platform at 860 nm [J]. IEEE Journal of Quantum Electronics, 2020, 56(1): 1–8.
165. Hu H, Fang Q, Wang S, et al. Folded 15-channel arrayed waveguide gratings on silicon nitride platform with loop reflector-Science Direct [J]. Optics Communications, 2021, 500(21): 1–8.
166. Taha, Ayat M., et al. Compact MMI-Based AWGs in a Scalable Monolithic Silicon Photonics Platform [J]. IEEE Photonics Journal, 2021, 13(4): 1–6.
167. Liu Guangcan, Bai Yanzhu, Planar Optical Waveguide Devices and Application Analysis [J]. Application of Optoelectronic Technology, 2005, 20 (3): 45~48
168. Zhang Kecang, Wang Ximin, Nonlinear Optical Crystal Materials Science. Beijing: Science Press, 2005.
169. Hou Zhipei. Effects of applied electric field and temperature on the optical properties of lithium niobate devices [D]. Tianjin: Tianjin University, 2014.
170. Wang Jiaxing. Design and fabrication of silicon dioxide planar optical waveguide splitter [D]. Beijing: Beijing University of Technology, 2013.
171. Gong Xiyu, Zhang Peng, Wu Xiaojie, et al. Research on the influence of aberration and turbulence on the performance of 90° spatial optical mixer [J]. Chinese Journal of Photonics, 2021, 50 (4): 74–88.
172. Jiang Yinli. Research on terahertz wave modulation technology and terahertz wave filtering technology based on optical mixing devices [D]. Wuhan: Huazhong University of Science and Technology, 2011.
173. Liu Hongzhan, Ji Yuefeng, Xu Nan, et al. Influence of signal and local oscillator amplitude distribution on mixing efficiency of inter-satellite coherent optical communication system [J]. Chinese Journal of Optics. 2011, 31 (10): 1–6
174. Li Xiangyang, Ma Zongfeng, Shi Dele. Influence of light field distribution of Gaussian beams on mixing efficiency of coherent detection [J]. Infrared and Laser Engineering. 2015, 44 (2): 539–543
175. China Southern Airlines, Zhang Peng, Tong Shoufeng, et al. Analysis of the influence of light spot size deviation and optical axis deflection on the mixing efficiency of spatial optical mixer [J]. Infrared and Laser Engineering, 2017, 46 (4): 205–212.
176. Ma Tingting. Analysis of the influence of external factors on the mixing efficiency of spatial optical mixers [D]. Changchun University of Science and Technology, 2019.
177. Pei Aihua, Zhang Shuren. Research on balance detector in coherent optical fiber communication system [J]. Journal of Changchun Institute of Optics and Fine Mechanics, 1993, 16 (4): 12–18.
178. Tan Zhenkun, Ke Xizheng. Mixing efficiency in coherent detection systems [J], Advances in Laser and Optoelectronics, 2017, 54:100605 (9).
179. David Fink, Coherent Detection Signal-to-Noise. APPLIED OPTICS. 1975, 14(3): 689–690

180. Wang Chao, Jiang Lun, Dong Keyan, et al. Analysis of the polarization characteristic of a satellite-to-ground laser communication optical system [J]. Lasers and Optoelectronics Progress, 2015, 52(12):120–126.
181. S A Maas, Nonlinear Microwave and RF Circuits. Second edition [M]. Norwood, MA: Artech House, 2003.
182. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress of adaptive optics technology for Optical Wireless Communication systems in Xi'an University of Technology [J]. Strong Laser and Particle Beam, 2021, 33 (08): 30–52.

Principles and Development of Optical Amplifiers

30

Contents

Introduction 1130
Optical-Amplifier Classification 1130
Principle of Optical Amplifiers 1131
Erbium-Doped Optical-Fiber Amplifiers (EDFA) 1131
Principle of the Erbium-Doped Fiber Amplifier 1132
Research Progress on Erbium-Doped Fiber Amplifiers 1134
Raman Fiber Amplifiers (RFA) 1136
Principle of the Raman Amplifier 1137
Research Progress on Raman Amplifiers 1138
Semiconductor Optical Amplifiers (SOAs) 1140
Principle of Semiconductor Optical Amplifiers 1141
Classification of Semiconductor Optical Amplifiers 1141
Characteristics of Semiconductor Optical Amplifiers 1143
Research Progress on Semiconductor Optical Amplifiers 1144
Summary and Outlook 1146
Summary 1146
Outlook 1146
References 1147

Abstract

The working performance of an optical communication system is not only related to the light source, but also to its transmission medium. With the development of fiber optic technology and the pursuit of long-distance transmission in the field of optical communication, optical amplifiers have emerged. Optical amplifiers can directly amplify optical signals and have great application value in the field of communication. The basic principle and development of optical amplifier are reviewed in this paper. Firstly, the concept, classification types of optical amplifiers are introduced, including their working principles and performance. Then, the current situation of various types of optical amplifiers at home and abroad are analyzed. Finally,

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_30

the future development of optics and principle of optical amplifiers are introduced. Then, the key technologies of the three amplifiers are prospected.

Keywords

Erbium-doped fiber amplifier · Raman fiber amplifier · Semiconductor optical amplifier · Infinite optical communication

Introduction

Optical amplifiers play an important role in optical communications by amplifying the signal light so that it can be adapted to various applications. In the field of fiber-optic communication, the continuous advancement in fiber-optic fabrication technology has led to a very low level of intrinsic loss in fiber-optic transmissions. However, the insertion loss of fiber-optic connectors and the intrinsic loss of long-distance transmissions lead to the attenuation of optical signals, which seriously restricts the transmission distance of fiber-optic communication systems. Therefore, it is important to study methods for amplifying optical signals without electrical relays.

Channel attenuation causes the optical signal to attenuate to an undetectable level after a certain transmission distance. Before optical amplifiers were invented, photoelectric and electro-optical conversions were typically used to amplify the optical signal. That is, after the attenuation of the optical signal was detected, it was converted to an electrical signal, shaped and amplified, reloaded to the optical wave, and allowed to continue to propagate.

The photoelectric conversion often decreases the signal-to-noise ratio. Optical amplifiers can directly amplify the attenuated signal light without the optical-electrical-optical conversion [1]. As fiber-optic communication technologies have developed towards long distances, high speeds, and large capacities, optical amplifiers have become an important development direction for future fiber-optic communication technologies.

Optical-Amplifier Classification

There are two types of optical amplifiers: fiber and semiconductor. Fiber amplifiers can be classified into various types, including erbium-doped, praseodymium-doped, and Raman amplifiers. In the 1550 nm wavelength range, the gain of the erbium-doped optical amplifier is the best, whereas in the 1310 nm wavelength range, the gain of the praseodymium-doped optical amplifier is the best. Raman amplifiers are typically used for distributed amplification.

Semiconductor optical amplifiers have received considerable attention because of their small size and easy integration; however, the polarization effect limits their commercial application.

Principle of Optical Amplifiers

An optical amplifier amplifies optical signals using the effects of excited radiation and optical amplification. The operating principle of an optical amplifier is similar to that of a semiconductor laser, in that both use a constant carrier leap to produce excited emissions to achieve a gain effect. Hence, the optical amplifier can be considered as a laser without an optical resonant cavity or an amplifier with feedback suppression from the resonant cavity [2].

In an optical amplifier, the optical signal is transmitted to the amplifier through an optical fiber or other transmission medium, processed by a series of optical components, and then enters the amplification medium. The amplifying medium is typically an optical fiber or semiconductor material doped with rare-earth ions (such as erbium, europium, neodymium, or terbium).

When the optical signal passes through the amplifying medium, the excited-radiation effect gives the photons emitted by the rare-earth ions in the amplifying medium the same frequency as the optical signal. As the emitted photons interact with the photons of the original signal, they amplify the optical signal and produce a gain effect [3]. In addition, the photons in the amplifying medium are affected by the light-amplification effect, thus increasing the photon energy and amplifying the light signal.

Specifically, optical amplification can be divided into three steps: first, the input optical signal is guided into a working medium doped with an exciter, such as an optical fiber or semiconductor. Next, the exciter converts the input optical signal into a set of photons with a photon energy greater than that of the input photons. Finally, these newly generated photons interact with the original photons, thus amplifying the input optical signal. The amplification effect of the optical amplifier is related to the dopant concentration, length of the fiber or semiconductor, wavelength, and other factors.

In summary, the principle of an optical amplifier is to amplify an optical signal using excited radiation and the light-amplification effect. Optical amplifiers are widely used in the fields of optical communications and sensing.

Erbium-Doped Optical-Fiber Amplifiers (EDFA)

Erbium-doped fiber amplifiers were the first commercially available optical amplifiers [4]. They are widely used and simple to construct [5]. They use a wavelength-selective coupler to mix a higher-power beam with an input signal that is significantly different in wavelength from the excitation light, and direct the mixed light into an erbium-containing fiber [6]. This high-energy beam stream is used to excite Er^{3+} to higher energy levels. When a signal with a wavelength different from that of the pump light collides with an excited Er atom, the Er atom loses energy and returns to a lower energy state [7]. An erbium-doped amplifier is a high-gain amplifier [8] that has the advantages of a high gain, wide bandwidth, low noise, and high output power.

Principle of the Erbium-Doped Fiber Amplifier

Erbium-doped fiber is formed by doping the core of a quartz fiber with the trivalent rare-earth metal erbium; Er^{3+} ions leap to higher energy levels under the action of a pump light. The radiation excited by the action of the signal light drops to the ground state and emits photons corresponding to the signal light as it does so [9]. This increases the number of photons, thus amplifying the signal [10]. The erbium-doped fiber-amplifier pumping process uses a three-energy-stage structure, the principle of which is shown in Fig. 30.1.

By injecting an extremely strong pumping light into the Er-doped fiber, more Er^{3+} can be transported from the ground state to the excited state, thus efficiently pumping the Er^{3+}. On this basis, electrons are rapidly transported from the excited state to the substable state. Because the erbium ion has a longer lifetime at a substable energy level, particle inversion can easily occur in this case.

In this process, the signal photon interacts with the doped rare-earth ions after passing through the doped fiber, resulting in excited radiation. This creates photons corresponding to the signal light on its surface, which leads to a sharp increase in the number of signal photons propagating through the fiber, thereby acting as a signal amplifier. When the signal light is amplified, the particles in the substable state enter the ground state via amplified spontaneous emission (ASE) [11]. This ASE is also amplified, which reduces the pump light and introduces noise [12].

When the power of the signal light and amplified spontaneous emission is low, the number of particles accumulated in the substable state is sufficient, and the gain can be maintained at a high level. When the pump-light supplement of substable particles

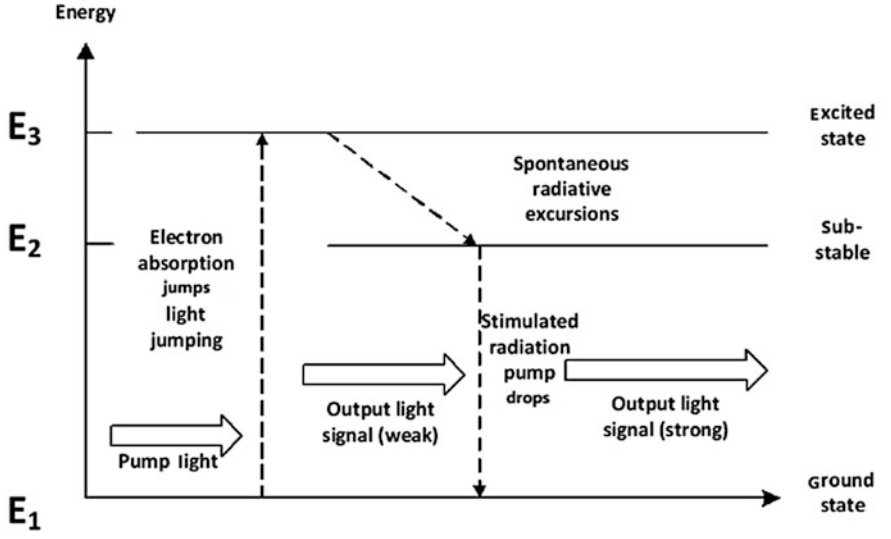


Fig. 30.1 Pumping process of an erbium-doped fiber amplifier

is insufficient to support the consumption of signal light and amplified spontaneous emission, the gain of the erbium-doped fiber amplifier drops and enters the amplification saturation state. Therefore, the suppression of ASE noise is an important direction in research on erbium-doped fiber amplifiers.

To amplify optical power, certain passive optical devices, pump sources, and erbium-doped fibers are combined into a specific optical structure to form an erbium-doped fiber amplifier. Figure 30.2 shows the optical structure of a typical erbium-doped optical fiber amplifier with a dual-pump source [12].

As shown in Fig. 30.2, the input-signal light is divided into two optical signals by the optical splitter. One goes to the photodiode (PD) light detector; the other passes through the fiber-optic isolator (ISO). It is then combined with the pump light emitted by the pump laser by the optical coupling element, and then passes into the erbium-doped fiber amplifier, which consists of two pump lights and two secondary pumps. Under the excitation of the pump light, the erbium-doped fiber amplifier amplifies the signal light, which then passes through the optical isolator and optical splitter and is output.

Features of Erbium-Doped Fiber Amplifiers

For erbium-doped fiber amplifiers, the amplifier gain is an important factor affecting the output performance; the amplification effect depends on the gain characteristics. The gain G of the erbium-doped fiber amplifier is $G = 10\ lg(P_o/P_i)$, which is the ratio of the output-signal optical power P_o to the input-signal optical power P_i.

The erbium-doping concentration in the fiber also significantly influences the gain. At an erbium-ion concentration greater than a certain value, the gain significantly decreases because too many erbium ions will polymerize and lead to concentration extinction; therefore, the concentration of doped erbium ions must be strictly controlled when preparing an erbium-doped optical fiber.

In addition to the concentration of erbium ions, the signal gain is closely related to the optical power of the pump. The relationship between the two is shown in Fig. 30.3. The gain coefficient of the system is greater for a smaller

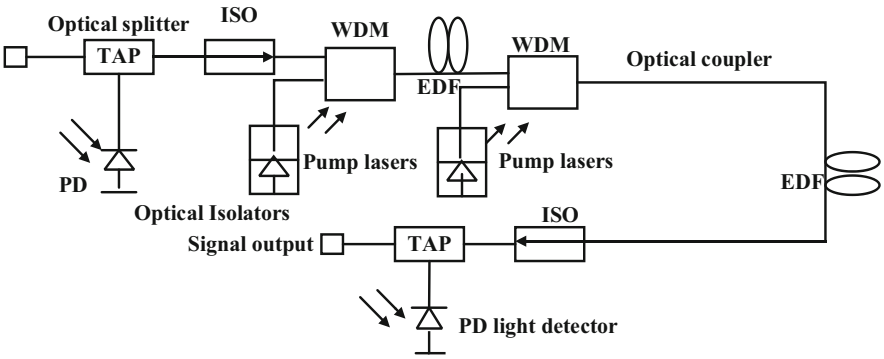


Fig. 30.2 Optical structure of an erbium-doped optical fiber amplifier [13]

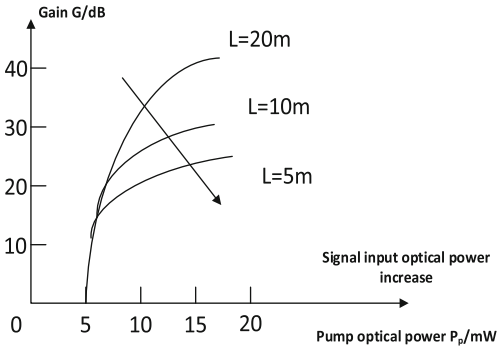


Fig. 30.3 Gain versus pump optical power [14]

input signal than for a larger input signal. Generally, when the gain is equal to 0, the optical power of the pump is called the threshold power P_{th}, whose expression is $P_{\text{th}} = AhP_p/BT$, where A is the effective area of action, h is Planck's constant, P_p is the pump optical power; B is the pump-absorption cross section; and T is the fluorescence lifetime (10^{-14} ms).

P_{th} is usually 5 mW. If $P_p = P_{\text{th}}$, the net gain is zero. When P_p continues to increase, G increases. The results show that in the initial stage, G increases rapidly and linearly with P_p. As P_p changes, G gradually decreases and eventually reaches the saturation region. When the pump optical power P_p has a ratio P_p/P_{th} greater than three, the gain of the amplifier tends to saturate [14]; that is, despite the increase in the pump optical power, the gain of the amplifier remains almost unchanged.

The gain of an erbium-doped fiber amplifier is also related to the length of the fiber. As the length of the Er-doped fiber increases, the gain tends to increase during the initial stage but decreases as the fiber length increases [15]. However, it should be noted that this length is only the length required to obtain the maximum gain, owing to other performance factors (e.g., noise performance).

Research Progress on Erbium-Doped Fiber Amplifiers

Research on doped fiber amplifiers began in 1962; however, owing to technological limitations, they were never implemented. In 1985, Mearas et al. introduced erbium-doped fiber-amplification technology, which began to develop rapidly and became a new research hotspot [16].

In 2012, Kang et al. from the University of Southampton, UK, theoretically demonstrated the characteristics of a step-refractive multimode erbium-doped fiber amplifier and simulated its single-mode gain of 20 dB and an intermodal-gain difference of 2 dB [17].

In 2013, Mahad et al. investigated gain flattening and obtained a gain flatness of 24 ± 0.29 dB in the range of 1546 nm–1558 nm [18].

In 2013, Jung et al. conducted an experimental study on the characteristics of a ring erbium-doped fiber. A two-signal mode gain was achieved in the experiment, and an average gain of 22 dB was obtained for a 3 dB intermodal-gain difference [19].

In 2014, Herbster et al. from the University of São Paulo, Brazil, chose a 180 mW LP11 pump light as the pump source and developed a new dual-ring Er-doped fiber. It obtained an average gain of 24.3 dB for four signal modes and a 0.6 dB intermodal-gain difference [20].

In 2015, Lecocq et al. proposed a novel fiber-design method based on a decreasing optimization method. They proposed a gain-equalization design method for microstructure-based erbium-doped fibers with few modes [21].

In 2015, Zhang studied the gain and noise of erbium-doped fiber amplifiers, based on the fiber mode-field theory, and reconfigured the spatial distribution of Er by adjusting the optical pumping structure to achieve a gain balance [22].

In 2016, Liu et al. conducted an experimental study on laser-amplification technology based on flat-plate waveguide lasers using flat-plate waveguide theory and laser-theory rate-dependent equations for flat-plate waveguide lasers [23].

In 2016, Yao et al. developed a novel erbium–ytterbium co-doped fiber amplifier using an auxiliary-cavity pumping technique and experimentally analyzed its spectrum, gain, and other performance metrics. A 17 W pump power was obtained in the 976 nm band, and 7.25 W was achieved in the 1550 nm band with a gain flatness of 0.12 dB for 100 mW of operation [24].

In 2017, Varona et al. constructed an erbium-ytterbium co-doped optical amplifier with continuous characteristics without amplified spontaneous emission, based on a main oscillating power-amplification structure. They replaced a 976 nm semiconductor laser with a 940 nm semiconductor laser, and obtained the highest laser-output power of 111 W at the central wavelength of the 1556 nm band with a 46.2% conversion efficiency [25].

In 2017, Savelii et al. of the *centre national de la recherche scientifique* (CNRS) optimized the glass composition of Er-doped alumina fibers, with Er ions in the fiber core layer, to incorporate a glass modifier [26].

In 2017, Liu proposed a novel low-noise and high-gain dual-way erbium-doped fiber amplifier, based on the theory of Er-ion excited amplification, and optimized its performance by numerical simulation [27].

In 2018, Zhang theoretically modelled a multiwavelength erbium-doped fiber amplifier based on the laser velocity equation and investigated its amplification performance and gain-balance performance [28].

In 2020, Lei et al. used values from a C-band light source to pump quartz fibers, with improved pumping efficiency compared to the conventional 1480 nm or 980 nm pumping modes, thus increasing the gain value in the L-band [29].

In 2021, Chen et al. from the Huazhong University of Science and Technology developed a Yb/Er co-doped fiber to improve its pumping efficiency by increasing the phosphorus content in the core to increase its maximum soliton energy. The core diameter of this fiber was 10 mm, the cladding diameter was 130 mm, and cladding absorption coefficient at 940 nm was 3.58 dB/m. The core absorption coefficient at

1535 nm was 34.5 dB/m. The laser slope efficiency at a 1.5 mm wavelength reached a maximum of 35.5%, with an output power of 2.5 W [30].

In 2021, Khudyakov et al. developed and tested a tapered erbium-doped fiber amplifier to achieve pulsed amplification with a high peak power. The core diameter was changed from 15.8 μm to 93 μm with a pump-to-signal conversion efficiency of 15.6% with a 1555 nm single-mode performance slope maintained along 3.7 m [31].

In 2021, Chen et al. developed a bidirectional erbium-doped fiber amplifier with low gain and low noise characteristics, based on a linear cavity structure, and obtained a gain factor of 15.52 dB and a noise factor of 3.25 dB [32].

In 2022, He et al. from the Huazhong University of Science and Technology successfully developed a new L-band quartz-based doped fiber using a modified chemical-vapor deposition method. They studied the effect of co-doped ions on the excited-state absorption characteristics of the material at the 4I13/2–4I9/2 energy level to establish single-stage forward-pumped and multi-stage optical-amplifier devices to study the broadband-amplification characteristics of the material.

Based on the forward 980 nm single-stage pumping structure, the gain reached 10.5 dB and the noise index reached 5.9 dB at a 9 dBm input signal power and a 530 mW pumping power at a 1629.3 nm wavelength. Using a multi-stage amplification structure, the gain of the fiber reached 27.4 dB at a 1629.3 nm wavelength [33].

Raman Fiber Amplifiers (RFA)

The structure of the Raman amplifier is similar to that of the erbium-doped fiber amplifier; however, because of the difference in the amplification medium, the output signal is amplified using the principle of Raman amplification. In contrast to other devices, such as erbium-doped fiber amplifiers and semiconductor optical amplifiers, the amplification of this device is based on a nonlinear interaction between the signal light and pump light. Raman fiber amplifiers use optical fibers, which have a simple structure, for nonlinear amplification [34].

A Raman amplifier is a fiber-optic amplifier that amplifies optical signals using the excited Raman-scattering effect [35]. In a nonlinear optical medium, the signal light at low frequencies interacts with the pump light at high frequencies to form a new signal light, and the excess energy is absorbed by the material in the form of molecular vibrations (i.e., photo-phonons), resulting in a jump between the two vibrational states. This process allows for the amplification of all-optical signals in the 1260–1675 nm range [36].

The pumped light can enter the transmission fiber along the same direction (isotropic pumping), in the opposite direction (reverse pumping), or be coupled in both directions (bidirectional pumping) [37], which yields an increased gain bandwidth of up to 6 THz [38]. Among these methods, reverse pumping is most common because of the reduced noise-transmission from the pump to the signal. These characteristics allow Raman amplifiers to be used in high-speed all-optical optical-communication systems.

Principle of the Raman Amplifier

A Raman fiber amplifier is an optical-signal amplification device based on excited Raman scattering in quartz fibers. It is achieved by simultaneously transmitting a weak signal and a strong pump light in the fiber in the Raman gain range of the pump light and amplifying the weak signal using the field of the excited Raman scattering effect.

Figure 30.4 illustrates the working principle of the Raman amplifier. The pump light is used to excite the electrons in the fiber so that they jump from low- to high-energy levels. On this basis, when the signal light is irradiated to its imaginary energy level, the electrons in its imaginary energy level emit Stokes photons to the higher-energy level of vibration. The rest of the energy is absorbed by the matter in the form of photo-phonons, causing the two vibrational dynamics to jump.

According to the law of conservation of energy, the internal energy of the molecules increases. This energy-transfer process is known as the Raman light-amplification process. A high-frequency pump light produces nonlinear scattering in the transmission medium, such that its incident light produces a Stokes frequency shift. The size of the shift depends on the vibrational energy level of the molecule, the amplitude of which determines the frequency range of the excited Raman scattering.

The Raman amplifier has a wide gain bandwidth, and with an increase in pump power, the amplifier's Raman gain also increases. However, the pump power cannot be increased indefinitely; if the pump power is transferred to the signal light, the resulting loss is large and will result in a gain saturation [39].

Features of the Raman Amplifier

The Raman amplifier has the following characteristics:

1. The gain wavelength is related to the pump wavelength. Theoretically, the signal can be amplified in any wavelength range when an appropriate pump source is selected. Using this method, a gain bandwidth can be obtained from multiple

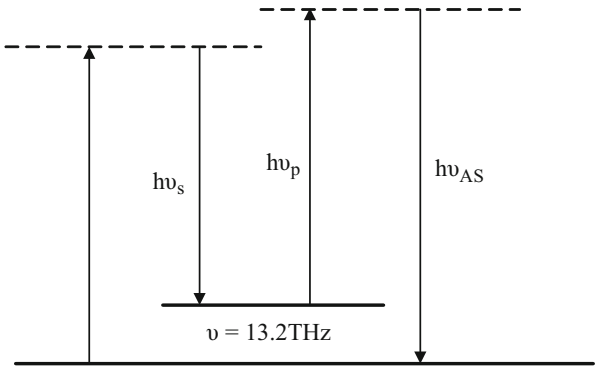


Fig. 30.4 Raman energy-level diagram of fiber-optic molecules

pump-light sources that is much larger than that of an erbium-doped fiber amplifier.
2. No other amplification medium is required because the gain medium is the transmission fiber itself. Thus, for those who find setting up repeaters inconvenient, fiber-optic communication systems have good application prospects.
3. Good noise performance. The signal power in the fiber is relatively small, regardless of the location. For ultralong-distance transmissions, Raman amplifiers are typically used in high-digit systems over kilometers to ensure a good signal-to-noise ratio.

Research Progress on Raman Amplifiers

In 1928, Raman experimentally discovered Raman scattering, where light waves changed their frequency when subjected to Raman scattering. Since this discovery, scientists have been studying Raman scattering [40].

In the 1960s, Koester of the United States suggested that rubidium-doped optical fibers could be used as optical-amplifier materials.

In 1972, Stolen et al. observed excited Raman scattering and the excited Raman effect in optical fibers and provided experimental results. Since then, scientists worldwide have studied optical-fiber communications based on excited Raman scattering [41]. However, erbium-doped fiber amplifiers received widespread attention because they accelerated the practical application of wavelength-division multiplexing (WDM) technology and adapted to the social needs of the time. After the 1990s, with the development of multichannel optical-amplification technology, erbium-doped fiber amplifiers were gradually restricted, and Raman fiber amplifiers reemerged.

In 1985, Aoki et al. first applied Raman amplifiers to fiber-optic communication systems to enable the transmission and amplification of light signals [42].

In 2002, Perlin et al. proposed an effective amplifier design scheme to optimize the gain and gain bandwidth of a multi-wavelength pumped Raman fiber amplifier. Compared with unoptimized Raman amplifiers of milliwatt magnitude, the gain flatness of this scheme was greatly improved. A gain spectrum with an 83 nm bandwidth and 0.05 dB flatness was obtained [43].

In 2003, Mori et al. used a 250 m tellurium-based fiber with a high Raman-gain coefficient and a four-wavelength channel laser as the pump source to achieve a Raman-gain spectrum with a 160 nm bandwidth, and a gain of more than 10 dB via bidirectional pumping [44].

In 2004, Wang et al. of Southeast University, China, used a 60 km-long general-purpose single-mode fiber as the gain medium in an experiment, and obtained a 18.5 dB average switching gain and 1.5 dB flatness for each channel by controlling the pump light from five reverse-pump sources [45].

In 2005, Varshney et al. used a multi-clad photonic-crystal fiber (PCF) with a 5.2 km length and a single 1450 nm pump source with a 520 mW power to obtain a

Raman-gain spectrum with an average net gain of 19 dB and a gain flatness of ±1.2 dB in a 30-nm bandwidth [46].

In 2006, Yan et al. used a genetic algorithm to optimize the wavelength and power of a Raman amplifier in backward pumping mode and obtained five optimal pumping wavelengths and powers. A distributed Raman-amplification system used a dispersion-shifted fiber with a 100.8 km gain medium and seven lasers in five wavelength ranges as pump sources. It obtained Raman amplification with an average gain of 15.55 dB and 0.87 dB over a bandwidth of 1520–1610 nm [47].

In 2008, Sperti et al. used a three-wavelength backward-pumping structure to optimize the pumping power using a hybrid genetic algorithm (combining a genetic algorithm and a simplex method). They found that the maximum difference between the experiment and simulation for the three channels was only 0.15 dB, when the preset signal optical gain was 7 dB in the test [48].

In 2009, Castellani et al. developed a low-flatness, high-gain, and low-loss photonic-crystal fiber Raman amplifier. This amplifier obtained Raman amplification in the C-band with an average gain of 8.5 dB and an average gain of 0.71 dB, using only two low-power pumped lasers [49].

In 2014, Singh et al. used a genetic algorithm to optimize the Raman-amplifier pump power in a wavelength-division multiplexing system with 320 channels spaced at 25 GHz intervals, and showed that the gain flatness of the Raman amplifier was less than 0.5 dB [50].

In 2017, Zhang et al. of the Harbin Institute of Technology used a new photonic-crystal optical fiber, optimized its structure, and reduced the Raman-gain flatness to less than 1 dB using one pump source [51].

In 2020, Tan et al. used a field-programmable gate array (FPGA) to control an eight-bit digital-to-analog converter (DAC) chip to provide the drive signal for a Fabry-Perot fiber tunable filter. It scanned in the 1532 nm–1567 nm band, and fed the detection results to the control terminal, thus achieving a tunable laser output [52].

In 2021, Song et al. developed a Raman optical amplifier based on the Yb band using a random fiber laser. By optimizing the random-fiber laser linewidth, a 3-dB output with a spectral linewidth of only 0.23 nm at a 100-W output power was achieved [53].

In 2022, Yao et al. of the National University of Defense Technology (NUDT), China, obtained the highest-power Raman-laser output of 373 W using spectral-synthesis techniques to enhance the pump-source brightness through laser-diode (LD) direct-pump Raman-amplifier research. The output-signal optical power was 373 W at a pump power of 544 W. The remaining pump light was 157 W, corresponding to an efficiency of 67.2% for light–light conversion [54].

In 2022, Gong of the Xi'an University of Posts and Telecommunications used a tellurium-based fiber as the transmission medium. Based on this, a second-order Raman-fiber amplifier consisting of two second-order pumps and four first-order pumps was proposed, which could be used to amplify the signal light in the C + L band. The second-order tellurium-based Raman fiber amplifier had an average

output gain of 31.3601 dB and a gain flatness of 0.6601 dB at an ultrahigh bandwidth of 1530–1630 nm [55].

Semiconductor Optical Amplifiers (SOAs)

A semiconductor optical amplifier generates gain through a semiconductor medium. The structure of the semiconductor optical amplifier is illustrated in Fig. 30.5. This amplifier is similar in structure to a Fabry-Perot excitation diode but with an anti-reflection element at its end. Because this causes the cavity to consume more power than its gain, it does not allow the amplifier to function as a laser. Another type of semiconductor optical amplifier consists of two parts. One is a Fabry-Perot semiconductor laser, and the other is a cone-type laser for reducing the power density of the output surface.

Currently, semiconductor optical amplifiers are primarily manufactured from Group II-VI GaAs/AlGaAs, InP/InGaAs, InP/InGaAsP, InP/InAlGaAs, and other semiconductors with Group II-VI structures. The amplifiers are generally used in communication systems, which operate at a wavelength of 0.85–1.6 μm, and their gain can reach 30 dB.

Semiconductor optical amplifiers have the advantages of miniaturization, electrical pumping, lower cost, and integration with other devices. They have high noise, small gain, low polarization correlation, high nonlinearity, and a fast transient response. The biggest advantage of semiconductor optical amplifiers is that they have four types of nonlinear effects (cross-gain modulation, cross-phase modulation, four-wave mixing, and polarization rotation). Their nanosecond state life gives their gain a fast response to changes in pump-light or signal-light intensity and also produces phase changes.

In addition, SOAs can operate at low laser power. Although the nonlinear effect is currently the biggest challenge in the field of fiber-optic communication, it also

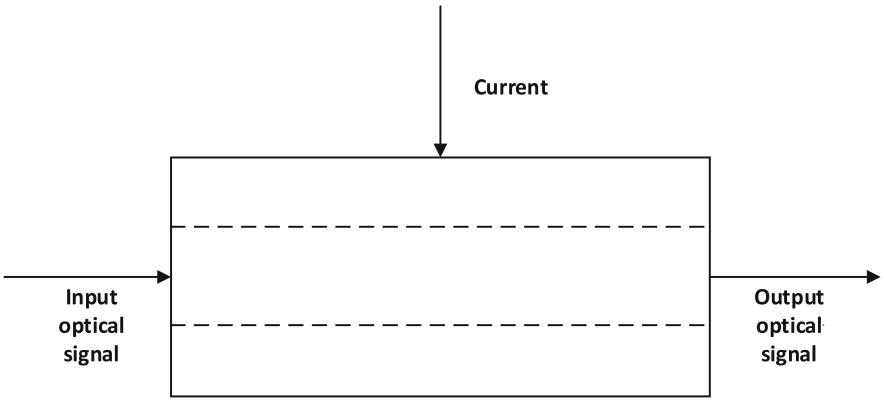


Fig. 30.5 Semiconductor optical-amplifier structure diagram [56]

offers the possibility of realizing wavelength bands different from those of erbium-doped fiber amplifiers. In recent years, some scholars have developed "linear optical amplifiers" using gain-clamping techniques.

Owing to their highly nonlinear characteristics, SOAs have great potential for applications in all-optical switching and wavelength conversion. Semiconductor optical amplifiers are core devices for optical signal processing, wavelength conversion, and other functions, and have a wide range of application prospects.

SOAs have a small size, high gain, and wide bandwidth, and can realize ultra-high-speed optical-pulse amplification at the picosecond level and output-signal amplification. SOAs have a variety of nonlinear characteristics that can convert signals between multiple wavelength bands in all-optical communication systems and can add multiple converters to optical networks. They are closely related to the development of wavelength-division multiplexing technology.

Principle of Semiconductor Optical Amplifiers

The difference between a semiconductor optical amplifier and a rare-earth-doped fiber amplifier is that the former's light-emitting medium is not a rare element but electron–hole pairs. The SOA is a new type of optical amplifier based on the mechanism of light amplification by excited radiation.

On this basis, a current simultaneously injects electrons and holes into the excitation region so that the electron–hole pairs in the excitation region produce a particle inversion, which results in a stronger light from the electron–hole pairs in the excitation region. When a semiconductor amplifier with light input power is considered, as shown in Fig. 30.6, the light input power is set to P_{in} and its output power to P_{out}; the output can be expressed as follows [57]:

$$P_{\text{out}} = P_{\text{in}} \times xp(agL), \tag{30.1}$$

where a denotes the proportion factor by which the light energy transmitted in the activation zone is suppressed [58].

Classification of Semiconductor Optical Amplifiers

Semiconductor optical amplifiers can be classified into two types. One is a normal semiconductor laser. The second uses an antireflection film on the two end faces of a Fabry–Perot (F-P) laser, which achieves broadband, high output, and low noise [59].

The multiple reflections generated by the semiconductor laser at the Fabry–Perot resonant-cavity junction are referred to as the F-P resonant cavity. As can be seen in Fig. 30.7, the incident light enters from the left end and passes through the active layer before reaching the right end. Some of the light is reflected back through the active layer, reaches the left end face, is amplified, and part of it exits from the left

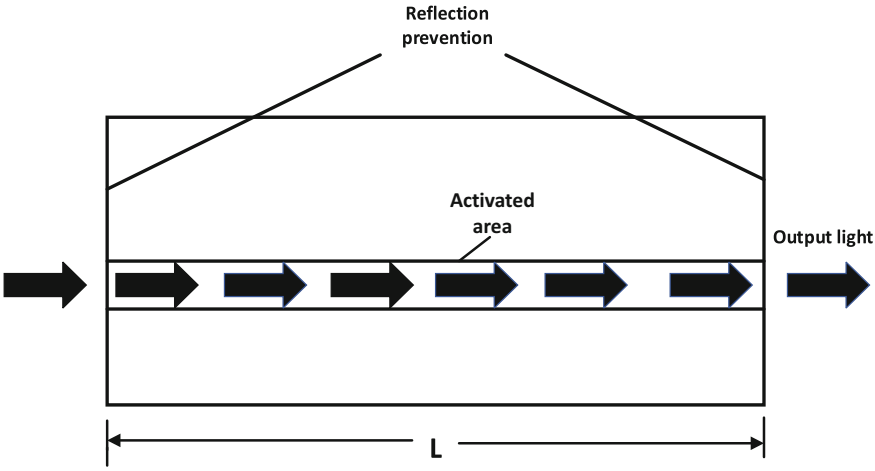


Fig. 30.6 Semiconductor optical amplifier

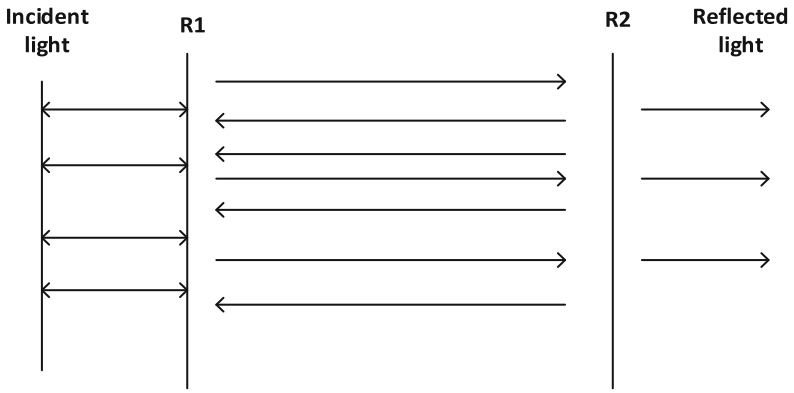


Fig. 30.7 Operating principle of an F-P semiconductor optical amplifier

end face. The remaining part is reflected from the left end face and passes through the active layer again, thus obtaining the amplification effect [60].

The device uses a Fabry–Perot resonant cavity to effectively modulate the optical signal. However, unlike semiconductor lasers, the light wave passing through a semiconductor optical amplifier is a traveling wave. This is because semiconductor optical amplifiers do not have resonant cavities or produce reflections. The operating principle of a traveling-wave semiconductor optical amplifier is shown in Fig. 30.8 [61].

In a semiconductor optical amplifier, the forward-injected current supplies carriers to the active region; when the injected current reaches a certain value, a particle-

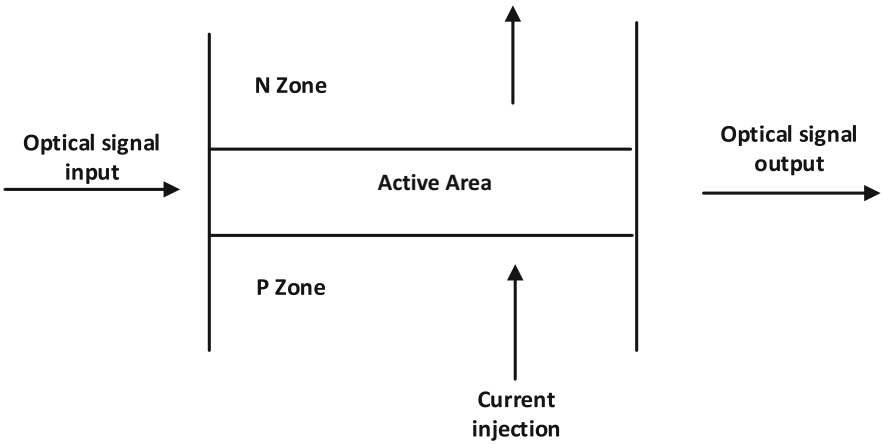


Fig. 30.8 Schematic diagram of a traveling-wave semiconductor optical amplifier

number flip occurs in the active region. If an optical signal is injected, the carriers in the semiconductor optical amplifier will be excited to produce photons. Because the light emitted by its excitation is the same as the surrounding light—that is, it has the same phase, polarization, and other information—the light can be amplified. The essence of semiconductor optical amplifiers is the interaction between photons and electrons in semiconductor materials. The higher the injected current, the higher the number of photons produced [62].

Characteristics of Semiconductor Optical Amplifiers

When the input signal is weak, the gain of a semiconductor optical amplifier is mainly manifested as amplified spontaneous emission; the power gain at this time is very low. When the input signal is sufficiently weak, its gain will rise sharply with the increase in current. At this time, the semiconductor optical amplifier works in the linear zone; i.e., its gain and the input signal are in a linear relationship. As the input signal increases, the gain flatness of the SOA gradually increases; however, the gain changes more slowly and approaches saturation.

In addition, the gain-saturation problem in semiconductor optical amplifiers has not attracted sufficient attention. When the input signal is constant, the gain of the semiconductor optical amplifier increases with the injection current; however, within a certain range, the gain decreases instead of increasing. This saturation is caused by the limitation of the carrier gain-recovery time [63].

The polarization characteristics of a semiconductor optical amplifier are primarily reflected in its sensitivity to the polarization state of the input signal light, and the gain of the signal light differs for different polarization states.

Research Progress on Semiconductor Optical Amplifiers

Semiconductor optical amplifiers have been investigated since the invention of semiconductor lasers in 1960 and the development of double-heterojunction devices in 1970 [64].

In 1970, Panish et al. prepared by solution epitaxy a double-heterostructure AlxGa1xAs–GaAs–AlxGa1xAs injected semiconductor laser with a room-temperature threshold as low as 2300 A/cm^2 [65].

In 1973, Zeidler et al. investigated the application of optical amplifiers in fiber-optic communication systems, where they could act as analog repeaters between regenerative repeaters in hybrid digital–optical communication systems [66].

In 1986, researchers began investigating InGaAsP semiconductor optical amplifiers because of the low loss of optical fibers at 1310 and 1550 nm. Broadband traveling-wave semiconductor optical amplifiers at 1500 nm were studied, analyzed, and further developed by Saitoh et al. [67].

In the 1990s, Mork et al. proposed a model for semiconductor optical amplifiers that included the local-carrier concentration and incorporated the temperature characteristics of carriers. This allowed better simulation of in-band effects, such as spectral hole-burning and free-carrier absorbance [68].

Since the 1990s, along with the continuous improvement of the semiconductor-technology level, research on semiconductor optical amplifiers has increasingly improved, especially with the progress of strain quantum wells and superlattice technology. Their application scope and performance were developed, especially in terms of the minimum signal gain, maximum output power, and signal-to-noise ratio index [69].

In 1990, Eisenstein et al. prepared a multi-quantum-well optical amplifier in the 1550 nm band and analyzed its size signal-gain characteristics [70].

In 2001, Connelly et al. established a steady-state model for semiconductor optical amplifiers, providing a broadband model of SOA gain. It was characterized by considering the relationship between the wavelength variation and gain [71].

In 2007, Liu et al. proposed the Hakki–Paoli direct-detection (HP-DD) algorithm, based on wavelet noise reduction and deconvolution. The results of this project helped to improve the measurement accuracy of the gain-coefficient spectrum of SOAs and reduce the influence of factors, such as instrument resolution and system noise, on the measurement accuracy. The simulated real spectra were obtained using a numerical-simulation process at resolutions of 0.05 nm, 0.1 nm, and 0.5 nm [72].

In 2011, Yu et al. conducted a study on all-optical logic devices based on cross-gain modulation for quantum-dot SOAs. Since quantum-dot semiconductor optical amplifiers have a fast gain-effect extraction property, their repair process can be accelerated by reducing the minority-carrier repair time, either by increasing the injected voltage or by decreasing the chirality period between the carriers from the wetted layer to the excited state [73].

In 2013, Ehsan et al. investigated the relationship between a tapered waveguide quantum-dot semiconductor optical amplifier and gain and cross-gain modulation;

the active region of a quantum-dot semiconductor optical amplifier has a tapered structure; therefore, the carrier-concentration distribution in the active region increases gradually with the direction of light propagation. Therefore, the gain and cross-gain modulation efficiency are improved [74].

In 2015, Matsumoto et al. used a strain-compensation method to prepare a 155 nm-band multilayer quantum-dot SOA to study the fundamental gain and femtosecond optical impulse response of this SOA. They laid the foundation for applying this semiconductor optical amplifier to ultrafast all-optical logic devices [75].

In 2016, Kotb et al. enabled all-optical logic with heterodyne gates, based on quantum-dot semiconductor optical amplifiers, to achieve an operating rate of 1 Tbit/s with a high output performance [76].

Recent Advances in Quantum-Dot Semiconductor Optical Amplifiers (QDSOA)

In 2017, Yang et al. conducted a theoretical and experimental study of cross-gain modulated/cross-phase modulated quantum-dot all-optical wavelength conversion, based on quantum-dot semiconductor amplifiers. They systematically investigated the photoelectric-conversion efficiency of cross-gain modulated/cross-phase modulated photonic-crystal wavelength-conversion devices by combining theory and experiments. Simulation results showed that reducing the input pulse width, decreasing the loss coefficient, shortening the leap time, increasing the length of the active region, and increasing the gain could improve the conversion efficiency of the wavelength converter [77].

In 2020, Mi et al. studied the cross-gain-modulated wavelength conversion and cross-phase-modulated wavelength conversion of quantum-dot semiconductor optical amplifiers. They conducted an in-depth study on the extinction ratio, Q-factor, and conversion efficiency of the method, and compared the method with existing cross-gain modulation and cross-phase modulation [78].

In 2022, Li et al. conducted a systematic study of the wavelength-conversion characteristics of photonic crystal-based quantum-dot SOAs under different external-field conditions using theoretical calculations, experimental tests, and numerical simulations. The results showed that the performance of photonic crystal-based quantum-dot SOAs was superior to that of quantum-dot SOAs [79].

Recent Advances in Reflective Semiconductor Optical Amplifiers (SOA)

In 2017, a new design idea for a gain-saturated reflective quantum-dot semiconductor optical amplifier was proposed by Yin et al. With appropriate parameter settings, the 3 dB cut-off frequency of the reflective quantum-dot semiconductor optical amplifier was expected to exceed 10 GHz [80].

In 2020, Deng et al. modelled and simulated a reflective semiconductor optical amplifier; when the bias current exceeded 130 mA, the gain of the reflective SOA remained above 26 dB and the noise figure remained at 11 dB. When the bias current was constant, the reflective SOA could still reach the optimal operating range at a low input power of −30 dBm [81].

In 2020, Rizou et al. experimentally demonstrated the feasibility of using birefringent fiber rings to enhance the efficiency of semiconductor optical amplification. A birefringent fiber ring is an optical filter that inverts the electro-optical characteristics of a reflective semiconductor optical amplifier. It can effectively eliminate the modulation bandwidth of a reflective SOA, limiting it in practical applications. By suitably adjusting its detuning, the birefringent fiber circuit can modulate the physical spectral composition, owing to the fiber detuning [82].

In 2021, Mukherjee et al. first used gain dynamics based on a reflective semiconductor optical amplifier to design and analyze a logical gate using Arcon pulses. The simulation results showed that the gate had efficient performance ($Q > 80$ dB) with low hardware complexity. The logic gate had an extinction ratio of approximately 14 dB, contrast of approximately 15 dB, and Q of approximately 90 dB. The effect of amplified spontaneous emission noise was also investigated. The logic gate used a single reflective semiconductor optical amplifier and can be used in the future design of any optical logic processor [83].

Summary and Outlook

Summary

Optical amplifiers have evolved rapidly over the past few decades and have become major components in today's communication networks. Optical amplifiers have attracted considerable attention in both academia and industry. They will become important components of future all-optical communication systems. With the continuous progress of optical wavelength-division multiplexing technology, the amount of data transmitted in optical fibers will become increasingly large, which means that the future development direction of optical-fiber communication is to achieve all-optical communication.

The invention of optical amplifiers, on this basis, has promoted the development of optical-fiber communication technology and organically combined it with the future development of new equipment and technology. It is expected to gradually develop in an all-optical direction.

Outlook

As one of the most important technologies in optical communication, optical amplifiers have a promising future development. Future development of erbium-doped fiber amplifiers will likely start from the following points.

1. Research on new highly doped fibers compatible with silicon-based fibers is important for high-power erbium-doped fiber amplifiers.
2. Research on new co-band high-power pump lasers is expected to be a key technical means of reducing the power output of erbium-doped fiber amplifiers.

3. Investigating new gain-flattening treatments is expected to broaden the range of bands amplified by erbium-doped fiber amplifiers in photoacoustic-spectroscopy detection systems.

Further developments in Raman fiber amplifiers could include the following:

1. Researchers should investigate the output gain characteristics of Raman fiber amplifiers with a high-order pumping structure.
2. Higher pumping-power mechanisms should be investigated to obtain higher-power pumping for Raman fiber amplifiers.

The future development of semiconductor optical amplifiers could start from the following points:

1. Research on the effects of the polarizability, end-face reflectivity, and carrier temperature of semiconductor optical amplifiers on wavelength conversion.
2. Research on monolithic integration technology for semiconductor optical amplifiers and other devices.
3. Research on reducing the optical-signal gain saturation.

References

1. Han Jingnan. Application of hybrid fiber amplifier in long-distance transmission [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2010.
2. Zhang Mingxiao. All-optical sampling technology of semiconductor optical amplifier in new generation fiber optic communication system [D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
3. Shang Zuguo. Design and fabrication of 600mW laser pump source control system [D]. Jilin: Jilin University, 2015.
4. Srivastava A K, Sun Y, Zyskind J L, et al. EDFA transient response to channel lossin WDM transmission system [J], Photonics Technology Letters, 2002, 14(3): 296–297.
5. Shen S, Naftaly M, Jha A. Tungsten-tellurite-a host glass for broadband EDFA [J]. optics Communications, 2002, 205(1): 101–105.
6. Harun S W, Poopalan P, Ahmad H. Gainen hance mentinL-band EDFA through adouble-pass technique [J].IEEE. Photonics Technology Letters, 2002, 14(3):296–297.
7. M. Yamada, A. Mori, K. Kobayashi, et al. Gain-flattened tellurite-based EDFA with a flat amplification bandwidthof 76 nm [J].IEEE Photonics Technology Letters, 2002, 10(9): 1244–1246.
8. Min B, Yoon H, Lee W J, et al. Coupled structure for wide-band EDFA with gain and noise figure improvements from C to L-band ASE injection [J].IEEE Photonics Technology Letters, 2000, 12(5):480–482.
9. Cui Xiaorong. Research on fiber amplifier based on DWDM system [D]. Jiaozuo, Henan: Henan University of Technology, 2009.
10. Agrawal G P. Nonlinear Fiber Optics (Second Edition) [M]. Boston: Mass Academic, 1995.471–524.
11. Liu Lanlan, Yang Wei, Wu Chongqing, et al. Amplfied Spontaneous Emission of erbium-doped fibers under arbitrary waveform pulse pumping [J]. Journal of Photonics, 2014, 43(12): 103–109.

12. Dou Qingying. Study of gain-controlled EDFA [D]. Tianjin: Nankai University, 2006.
13. Wang Chunnan. Analysis and research of fiber optic amplifier [D].Shanghai: Shanghai Jiao Tong University, 2011.
14. Wang Wei. Research on automatic power control of EDFA based on AGC technology [D]. Jilin: Changchun Univ ersity of Technology, 2010.
15. Sun Junqiang, Huang Dexiu, Li Zaigang. Optimal fiber length and gain characteristics of erbium-doped fib er amplifier [J]. Optical Communication Technology, 1994, 18(01):26–30.
16. Zhang Chao. Study of gain and loss of optical fiber amplifier by Huygens' principle [D]. Changchun, Jilin: Changchun University of Technology, 2021.
17. Jun Y, Alam S, Li Z, et al. First demonstration and detailed characterization of a multimode amplifier for s pace divi-sion multiplexed transmission systems [J]. Opt.Express, 2011, 19(26): 952–957.
18. Farah Diana Mahad, Abu Sahmah M, Sevia Mahdaliza Idrus, et al. Comparative performance testing of SOA wa velength conversion techniques for future all-optical systems [J]. Optik, 2013, 124(12): 1254–1259.
19. Jung Y, Kang Q, Sleiffer V, et al. Three mode Er^3+ ring-doped fiber amplifier for mode-division multiple xed transmission [J]. Optics express, 2013, 21(8): 10383–10392.
20. Zhao Qinghua. Study on gain equalization of erbium-doped fiber amplifier with few modes [D]. Tianjin: Tia njin University, 2018.
21. Cocq G L, Quiquempois Y, Bigot L. Optimization algorithm applied to the design of few-mode erbium doped fiber amplifier for modal and spectral gain equalization [J]. Journal of Lightwave Technology, 2015, 33(1): 100–108.
22. Zhang Zhenzhen. Simulation study of erbium-doped fiber amplifier with few modes [D]. Tianjin: Tianjin Univ ersity, 2015.
23. Liu Jiao. Research on high power planar waveguide laser amplification technology [D]. Mianyang, Sichuan, China: Chinese Academy of Engineering Physics, 2018.
24. Yao Yunchi. Research on high-power erbium-ytterbium co-doped fiber amplifier with auxiliary cavity pumpi ng [D]. Tianjin: Tianjin University, 2019.
25. Varona O D, Fittkau W, Booker P, et al. Single-frequency fiber amplifier at 1.5 m with 100 W inthe linearly- polarized TEM00 mode for next-generation gravitational wave detectors [J]. OpticsExpress, 2017, 25(21):2488 0.
26. Savelii I, Bigot L, Capoen B, et al. Benefit of Rare-Earth"Smart Doping"and Material Nano-structuring fo r the Next Generation of Er-Doped Fibers [J]. Nanoscale Research Letters, 2017, 12(1): 206.
27. Zang Qi, Deng Xue, Liu Jie, et al. Optimal design of bidirectional erbium-doped fiber amplifier for long -range optical frequency transfer links [J]. Journal of Optics, 2017, 37(03):197–204.
28. Zhang Bo, Zhang Entao, Hu Xiaochuan, et al. Study on the amplification characteristics of multi-wavelength erbium-doped fiber laser amplifier [J]. Laser Technology, 2018, 42(03): 325–330.
29. Lei C, Feng H, Messaddeq Y, et al. Investigation of C-band pumping for extended L-band EDFAs [J]. JOSA B, 2020, 37(8): 2345–2352.
30. Chen Yang, Chu Yingbo, Dai Nengli, et al. Preparation of erbium-ytterbium co-doped fiber and its laser p erformance [J]. China Laser, 2021, 48(07):47–53.
31. Khudyakov Maksim M, Levchenko Andrei E, Velmiskin Vladimir V, et al. Er-Doped Tapered Fiber Amplifier for H igh Peak Power Sub-ns Pulse Amplification [J].Photonics, 2021, 8(12):527.
32. Chen Wenjia, Cheng Chunfu, Ou Yiwen. Optimal design of low-gain and low-noise gain-clamped bidirectional erbium-doped fiber amplifier [J]. Optical Communication Technology, 2021, 45(10):14–18.
33. He Le, Chu Yingbo, Dai Nengli, et al. Quartz-based L-band extended erbium-doped fiber and its amplificat ion performance [J]. Journal of Physics, 2022, 71(15):143–149.
34. Bromage J. Raman Amplification for Fiber Communications Systems [J]. Lightwave Technology Journal of, 2003, 2 2(1):79–93.

35. M.D. Mermelstein, Headley C, Bouteiller J C, et al. Configurable three-wavelength Raman fiber laser for Raman amplification and dynamic gain flattening [J].IEEE Photonics Technology Letters, 2022, 13(12):1286–1288.
36. Nielsen T N, Stentz A J, Rottwitt K, et al. 3.28-Tb/s transmission over 3×100km of nonzero-dispersion fiber using dual C-and L-band distributed Raman amplification [J].IEEE Photonics Technology Letters, 2000, 12(8): 1079–1081.
37. Headley C, Agrawal G P. Raman Amplification in Fiber Optical Communication Systems [M]. Elsevier Academic Pr ess. 2005.
38. Rodriguez-Barrios F, Martin-Lopez S, Carrasco-Sanz A, et al. Distributed Brillouin Fiber Sensor Assisted by First-Order Rman Amplification [J]. Journal of Lightwave Technology, 2010, 28(15):2162–2172.
39. Dai Junke. Design and fabrication of wavelength division multiplexing light source detection module [D]. Hefei, Anhui, China: Hefei University of Technology, 2018.
40. Raman C V, Krishnan K S. A new type of secondary radiation [J]. Nature, 1928, 121(3048): 501–502.
41. Stolen RH, Ippen EP. Raman gain in glass optical wave guides [J]. Applied Physics Letters, 1973, 22(6):276–278.
42. Aoki Y, Kishida S, Washio K, et al. Bit error rate evaluation of optical signals amplified via stimulate d Raman process in an optical fibre [J]. Electronics Letters, 2007, 21(5):191–193.
43. Perlin V E, Winful H G. Optimal design of flat-gain wide-band fiber Raman amplifiers [J]. Journal of ligh twave technology, 2002, 20(2):250–254.
44. Mori A, Masuda H, Shikano K, et al. Ultra-wide-band tellurite-based fiber Raman amplifier [J]. Lightwave Technology, Journal of, 2003, 21(5): 1300–1306.
45. Wang Zhenyuan, Xia Lin, Wang Fang. Experimental study of C+L-band distributed Raman fiber amplifier [J]. Jou rnal of Southeast University (Natural Science Edition),2004, 34(03): 315–319.
46. Varshney S K, Saitoh K, Koshiba M. A novel design for dispersion compen sating photonic crystal fiber Ra man amplifier [J]. Photonics Technology Letters, IEEE, 2005, 17(10): 2062–2064.
47. Yan Fenfen, Wang Kuiru, Yu Chongxiu, et al. Gain-flattening multi-wavelength pumped broadband Raman fiber a mplifier [J]. Optical Precision Engineering, 2006, 14(2): 155–158.
48. Sperti D, André P S, Neto B, et al. Experimental assessment of some Raman fiber amplifiers solutions for coarse wavelength division multiplexing applications [J]. Photonic Network Communications, 2008, 16(3):19 5–202.
49. Castellani C E S, Cani S P N, Segatto M E, et al. Design methodology for multi-pumped discrete Raman amp lifiers: case-study employing photonic crystal fibers [J]. Optics express, 2009, 17(16): 14121–14131.
50. Singh S, Saini S, Kaur G, et al. On the optimization of Raman fiber amplifierus inggenetic algori thmin the scenario of a64nm320 channel sdensewave length division multiplex dsystem [J]. Journal ofthe Optical Socie ty of Korea, 2014, 18(2):118–127.
51. Zhang Hongyue. Gain equalization study of Raman fiber amplifier [D]. Harbin, Heilongjiang, China: Harbin Institute of Technology, 2015.
52. Tan Chao, Wan Shengpeng, Song Zaobiao, et al. Tunable laser based on FFP-TF2 filter [J]. Applied Optics, 2020, 41(01):186–193.
53. Song J, Lai W, Ma P, et al. Compact and low-cost superfluorescent-fiber-sourceassisted narrow linewid th Yb-Raman fiber amplifier [J]. Applied Optics, 2021, 60(6):1484–1488;
54. Yao Tianfu, Fan Chenchen, Xiao Hu, et al. LD-pumped Raman fiber amplifier achieves high brightness laser output for the first time [J]. Infrared and Laser Engineering, 2022, 51(06): 514–515.
55. Gong Jiamin, Jin Ku, Zhang Yi, et al. Optimized design of second-order tellurium-based fiber Raman ampli fier based on cooperative search algorithm [J]. Advances in Lasers and Optoelectronics, 2023, 60(19):1–16.

56. Xu Guiyong. Theoretical study of gain saturation characteristics of semiconductor optical amplifiers and wavelength conversion techniques [D]. Wuxi, Jiangsu, China: Jiangnan University, 2024.
57. Deng Canran, Zhu Yiqing, Xu Guiyong, et al. Modeling and simulation of reflective semiconductor optical amplifiers [J]. Applied Lasers, 2020, 40(03):532–538.
58. Jiang Tao, Chen Yan. Semiconductor optical amplifiers [J]. Advances in Laser and Optoelectronics, 2000, 337(8):40–45.
59. Yang Xianglin. Optical amplifiers and their applications [M]. Beijing: Electronic Industry Press.
60. Feng Xiancheng, Duan Wenxiao, Li Han. Characterization of gain semiconductor optical amplifiers [J]. Inte nse Laser and Particle Beam, 2012, 24(06):1475–1478.
61. Ding Keke, Ma Yuhang, Wei Long, et al. Advances in semiconductor optical amplifier research [J]. Optoelectro nics, 2021, 11(1):8–18.
62. Wu Chongqing. Optical-optical interactions in semiconductor optical amplifiers and their applications [J]. Physics, 2007, 36(8):631–636.
63. Chen Xu, Hu Zhegang, Xu Shengxiang. Gain characteristics of integrated dual-waveguide quantum well struc tured semiconductor optical amplifiers [J]. Electronic World, 2020,(09): 27–33.
64. Hayashi, I. Junction Lasers Which Operate Continuously Atroom Temperature [J]. Applied Physics Letters, 1970, 17(3):109–111.
65. Panish, B. M. Double-heterostructure injection lasers with room-temperature thresholds as low as 2300A/c m2[J]. Applied Physics Letters, 1970, 16(8): 326–331.
66. Zeidler G, Schicketanz D. Use of laser amplifiers in a glass-fibre communications system [J]. Radio & Ele ctronic Engineer, 1973, 43(11):675–682.
67. Saitoh T, Mukai T. 1.5. mu. m GaInAsP traveling-wave semiconductor laser amplifier [J]. IEEE J. Quant. Ele ctron, 1987, 23(6):1010–1023.
68. Uskov A, Mørk J, Mark J. Wave mixing in semiconductor laser amplifiers due to carrier heating and spectr al-hole burning [J]. IEEE Journalof Quantum Electronics, 1994, 30(8): 1769–1781.
69. Liu Zhiyong. Research on fiber laser based on semiconductor optical amplifier [D]. Jingzhou, Hubei, Chin a: Changjiang University, 2019:1–2.
70. Eisenstein G, Koren U, Raybon G, et al. Large- and small-signal gain characteristics of 1.5μm multiple quantum well optical amplifiers [J]. Applied Physics Letters, 1990, 56(13): 1201–1203.
71. Connelly M J. Wideband semiconductor optical amplifier steady-state numerical model [J]. IEEE Journal of Quantum Electronics, 2001, 37(3): 439–447.
72. Liu Lei, Zhang Xinliang, Huang Dexiu. Wavelet denoising and deconvolution based gain coefficient spectru m measurement method for semiconductor optical amplifiers [J]. China Laser, 2007, 34(6):846–850.
73. Yu Jianxing. Research on all-optical logic gate based on cross-gain modulation effect of quantum dot sem iconductor optical amplifier [D]. Chengdu, Sichuan: Southwest Jiaotong University, 2011:8–18.
74. E. Mohadesrad, K. Abedi. Design and Analysis of Multi-Electrode Tapered Quantum-Dot Semiconductor Optical A mplifiers [J]. Journal of Computational & Theoretical Nanoscience, 2013, 10(11): 2652–2662.
75. A. Matsumoto, Y. Takei, A. Matsushita, et al. Gaincharacteristics and femto-second optical pulse response of 1550 nm-band multi-stacked QD-SOA grown on InP(311)B substrate [J]. Optics Communications, 2015, 344(01):51–54.
76. A. Kotb. Simulation of high quality factor all-optical logic gates based on quantum-dot semiconductor opti cal amplifier at 1Tb/s [J].Optik, 2016, 127(1):320–329.
77. Yang Wenhua, Wang Hailong, Wang Zhaoxiang, et al. Wavelength conversion efficiency of quantum dot semico nductor optical amplifiers [J]. Journal of Optics, 2017, 37(04):94–101.
78. Mi Juchao. Study of quantum dot SOA-based cascaded XGM and XPM wavelength conversion [D]. Jining, Shando ng: Qufu Normal University, 2020: 54–70.

79. Li Xiaobo, Wang Hailong, Ma Linan, et al. Wavelength conversion characteristics of quantum dot semicondu ctor optical amplifiers based on photonic crystals [J]. Journal of Optics, 2022, 42(02):9–18.
80. Yin Yu, Ling Yun, Li Hao, Du Xiaojun, et al. High-pass characteristics of gain-saturated reflective semi conductor optical amplifiers with quantum dots [J]. Journal of Photonics, 2018, 47(01):19–30.
81. Deng Chanran, Zhu Yiqing, Xu Guiyong, et al. Modeling and simulation of reflective semiconductor optical amplifier [J]. Applied Lasers, 2020, 40(03):532–538.
82. Rizou Z V, Zoiros K E, Rampone T, et al. Reflective Semiconductor Optical Amplifier Direct Modulation Capability Enhancement Using Birefringent Fiber Loop [J]. Applied Sciences, 2020, 10(15):5333.
83. Mukherjee K,Maji K, Raja A, et al. All-optical soliton based universal logic NOR utilizing a single reflective semiconductor optical amplifier (RSOA) [J]. Photonic Network Communications, 2022, 43(2):101–108.

Key Technologies in Underwater Optical Wireless Communication

31

Contents

Introduction 1154
Underwater Wireless Communication 1154
Acoustic Communication 1154
Radio Frequency Communication 1156
Optical Communication 1159
Model and Research Status of Underwater Optical Wireless Communication Systems 1159
System Model 1159
Foreign Research Status 1161
Domestic Research Status 1167
Key Technologies in Underwater Optical Wireless Communication 1181
Channel Modeling of Underwater Optical Wireless Communication Systems 1181
Light Source Selection 1192
Modulation Technology 1193
Coding Technology 1193
Detection Technology 1194
Underwater Optical Wireless Communication System Design Model 1195
Summary 1201
References 1202

Abstract

Due to the advantages of high speed, low latency, and high confidentiality, underwater optical wireless communication (UOWC) has become the main means of underwater wireless communication. In recent years, this technology has developed rapidly, and the technical requirements for UOWC systems have also become increasingly high. This article provides a brief summary of the models and research status of the UOWC system, focusing on the key technologies used in the UOWC system, including the selection of light sources, modulation and coding, detection, channel modeling, and other key technologies. It also elaborates on the influencing factors and modeling of the theoretical channel model of UOWC. Finally, it summarizes the future development trends and challenges faced by UOWC for achieving long-distance underwater

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_31

communication and high-speed transmission. This provides a reference for the future development direction and practicality of UOWC.

Keywords

Underwater optical wireless communication · Modulation coding technology · Detection technology · Channel modeling

Introduction

Underwater optical wireless communication uses light as a carrier to transmit information. Its advantages include a high communication speed, small time delay, good security, and low cost. In recent years, researchers have conducted extensive research on key technologies in underwater optical wireless communication systems, including modulation and coding, detection, light source selection, and channel model establishment.

With the development of the Internet and the arrival of the 6G era, optical wireless communication combined with the Internet for ocean exploration is bound to be a future development trend. The study of key technologies in underwater communication will provide methods and technologies for future ocean exploration.

Underwater Wireless Communication

Ocean exploration requires real-time high-speed communication to transmit audio, image, and video information. Traditional communication uses wired technology, such as optical fibers and cables, to transmit information. However, cable communication is significantly affected by the underwater environment, and its construction difficulty and mobility are particularly poor. Moreover, it is very difficult to maintain and its cost is high.

Therefore, in recent years, wireless communication has been used in underwater communication systems to overcome the shortcomings of cable communication. Underwater wireless communication enables ocean exploration and is attracting increasing attention from researchers. Underwater wireless communication includes acoustic, radio frequency (RF), and optical communication [1].

Acoustic Communication

Sound waves are mechanical waves, and their attenuation in water is very small. Therefore, the transmission distance of sound waves in water is relatively long. Currently, acoustic wave technology for underwater wireless communication is more mature and also the most widely used.

Acoustic communication was first used in the military. In 1945, after the Second World War, the United States used sound waves to communicate in submarine systems; therefore, underwater acoustic communication rapidly developed [2]. Russia attached great importance to the development of aircraft carriers and invested considerable energy and financial resources in underwater submarine communication. The "God of the North Wind" submarine was Russia's military support [3].

With the continuous progress and development of science and technology, acoustic communication has been widely used in civil areas, such as monitoring underwater environmental pollution and underwater information collection. Digital signal processing technology has also been applied to underwater acoustic communication. Underwater acoustic communication uses sound waves as the communication carriers. After decades of development, the performance of underwater acoustic communication has stabilized [4].

Because the transmission speed of sound waves in seawater is related to the depth, temperature, salinity, etc., with increasing depth, the change in seawater temperature bends the sound waves. At a certain depth, upward divergent sound waves bend downward and the downward divergent sound waves bend upward. Therefore, sound waves oscillate back and forth near a certain depth, which is called the deep sea acoustic channel [4]. At this depth, acoustic energy transmissions are similar to laser transmissions in optical fiber, with minimal loss; therefore, they can be detected by sonar at a very long range. This is where the 550-km transmission distance was achieved by the United States.

The US Sound Surveillance System (SOSUS) sonar array on the seabed used the deep sea acoustic channel to detect enemy submarines thousands of kilometers away [5]. However, the depth in the South China Sea is approximately 1100 m, whereas in the Mediterranean and temperate Pacific, the depth of the acoustic axis is within the range of 100–300 m. In the Arctic Ocean, the acoustic axis is close to the sea [6]. However, the submarine is less than 1000 m deep; thus, its utilization value is very small [7].

Another underwater acoustic effect is that of the underwater acoustic convergence zone. Because sound waves bend underwater, waves emitted from shallow waters will focus at a distance of 35 km, which is called the first convergence area. They will focus again at 70 km, after being reflected by the sea surface; this is called the second convergence area. The third convergence area is at 105 km. Sound waves are very confused at this point, so there is generally no fourth convergence area [8].

The main problems of long-distance underwater acoustic communication are that the channel is complex and changeable, the noise interference is high, and the signal bandwidth is limited. Therefore, the communication rate is low [9]. In 2005, Northwest Polytechnical University conducted an underwater acoustic communication experiment with a 20-bit/s signal communication rate, 80 km from the coast [10]. In 2013, the joint research team of the US Navy and Marine Corps conducted an underwater acoustic communication experiment with a received signal communication rate of 6.5 bit/s at a distance of 550 km from the coast [9].

In addition, a team from the University of California, San Diego, United States, used orthogonal frequency-division multiplexing technology to achieve underwater

acoustic communication with a signal communication rate of 15 bit/s, 550 km from the coast, and 7.5 bit/s, 700 km from the coast [11]. Later, after improving the underwater communication system, the team finally conducted an underwater acoustic communication experiment with a signal communication rate of 50 bit/s at a location 550 km from the coast [12].

Acoustic communication is affected by underwater noise, salinity, and water pressure, among other factors. Most communication systems only use acoustic waves to transmit voice signals, which cannot adapt to complex battlefield conditions and are easily detected by enemy sonar detectors. Therefore, at present, underwater acoustic communication can only be used as an auxiliary communication means [13].

The underwater transmission time delay of acoustic waves is relatively large, and the communication rate can only reach the order of kbit/s [14]. In short-distance transmissions, the frequency of the acoustic wave is 50 Hz, and the communication rate can reach 100 kbit/s. In medium- and long-distance transmissions, the communication rate can reach 50 kbit/s when the acoustic frequency range is 50–100 Hz. For long-distance transmission, the communication rate can only reach 10 kbit/s [15] when the acoustic frequency range is 100–1000 Hz.

Underwater acoustic communication also has a large time delay, which is harmful to marine organisms [16]. Acoustic communication can easily be overheard and has poor confidentiality. The system equipment is bulky and consumes considerable power. Thus, it cannot meet the small size and low-power consumption requirements for underwater communication [17].

Radio Frequency Communication

The penetration ability of an RF signal in water is related to the RF signal frequency. The transmission distance of a 30–300-kHz signal in water is several meters, that of a 3 Hz–30 kHz signal in water is 10–20 m, and that of a 30–300-Hz signal in water is 100–200 m [17]. Currently, extremely low-frequency (ELF) and very-low-frequency (VLF) long-wave radio stations use RF signals for underwater communication. Relevant radio communication equipment has been established mainly in the United States, Russia, the United Kingdom, France, and other countries.

The US Navy built a number of long-wave radio stations for one-way underwater communication. In 1953, the Jim Creek VLF long-wave radio station was used for one-way underwater communication by Pacific Navy submarines. The radio call sign was NLK, the frequency was 28.8 kHz, and the output power was 1.2 MW [18]. In 1960, the Cutler VLF long-wave radio station in Maine was used for one-way communication between Atlantic submarines. The radio station call sign was NAA, the frequency was 24 kHz, and the output power was 1.8 MW [19].

In 1997, the LaMoure VLF long-wave radio station built in North Dakota was transferred to the US Navy for submarine communication. The radio station call sign was NML, and the frequency was 29.2 kHz [20]. In 1972, the Lualualei VLF long-wave radio station built in Hawaii served the US military in the Pacific theater, with

the call sign of NPM and frequencies of 24.4 kHz and 31.4 kHz [19]. The Harold E. Holt VLF long-wave radio station, located on the northwest coast of Australia, was used for surface and underwater ship–submarine communications in the Western Pacific Ocean and East Indian Ocean. The radio station was jointly used by the United States and Australia. Its call sign was NWC, and the frequency was 21.8 kHz [21].

The call sign of the Aguada, Puerto Rico, low-frequency long-wave radio station built by the US Navy was NAU, and the frequency was 40.75 kHz [22]. The call sign of the Keflavik, Iceland, low-frequency long-wave radio station was NRK, and the frequency was 37.5 kHz. The call sign of the Sigonella, Italy, low-frequency long-wave radio station was NSY, and the frequency was 45.9 kHz [23].

The VLF long-wave radio stations in Russia were mostly alpha navigation-base stations and beta time service stations. The call signs of the VLF long-wave radio stations used in the Russian navy were RSDN, RDL, and RJH, and the corresponding frequency parameters were 11.91 kHz, 31.2/24.1 kHz, and 25 kHz, respectively [23].

The British Anthorn long-wave radio communication station was located in Cumbria, England. Anthorn's VLF long-wave radio call sign was GQD, and the communication frequency to submarines was 21.6 kHz. Anthorn was one of the four VLF long-wave radio stations in the North Atlantic Treaty Organization (NATO). Anthorn's low-frequency long-wave radio is part of the British National Physical Laboratory. Its main function is to provide time service. The call sign of the radio is MSF with a 60-kHz frequency and 17-kW power [18]. In 2001, Skelton, England's very-low-frequency long-wave radio station was mainly used to send coding instructions to submarines. Its call sign was GBZ, and its frequency was 21.58 kHz [22].

France's very-low-frequency long-wave radio transmitter in Rosnay is an antenna system composed of 13 antenna poles. The central antenna pole is the highest at 357 m. The inner circle is composed of six 310-m-high antenna poles, and the outer circle is composed of six 270-m-high antenna poles. The call sign of the radio station is HWU, with frequencies of 15.1 kHz, 31.9 kHz, and 24.7 kHz [22].

The Norwegian Noviken VLF long-wave radio has a call sign of JXN and a frequency of 19.4 kHz. The Rhauderfehn VLF long-wave radio station in Germany has a call sign of DHO38 and a frequency of 31.4 kHz. These two types of long-wave radio stations are part of a submarine broadcasting system operated by NATO [24]. The Grimeton VLF long-wave radio in Sweden provides encrypted communication services for the Swedish Navy. The call sign of the radio station is SRC, and the frequency is 40.4 kHz. The radio station also has a channel with a call sign of SAQ and frequency of 19.2 kHz [25]. In 1976, Iceland's Grindavik VLF long-wave radio station was successfully built. In addition to VLF radio communication, stations can also achieve short-wave radio communication. The call sign of the station is TFK, and the frequency is 37.6 kHz [26].

The call sign of the Ebino VLF long-wave radio station in Miyazaki, Japan, is JJI, and the frequency is 31.2 kHz. The call sign of the Awase low-frequency long-wave radio station in Okinawa, Japan, is NDI with a 54-kHz frequency [26]. The call sign

of the Mokpo VLF long-wave radio station located in southwest South Korea is UNID25, with a frequency of 25 kHz [26].

India built VLF long-wave radio stations in Mumbai and south Tamil Nadu. The VLF long-wave radio station in Mumbai has been closed, whereas the VLF long-wave radio station in south Tamil Nadu is still in operation. The call sign of the radio station is VTX3, and the frequency is 18.2 kHz [25]. The Karachi VLF long-wave radio in Pakistan is used for submarine communication in the Pakistani Navy. The radio call sign is PNSH, and its frequency is 14–29.2 kHz [22]. Argentina's Monte Grande VLF long-wave radio is used for Argentine naval communications. The radio is composed of eight 210-m-high and two 219-m-high antenna poles. The call sign of the VLF long-wave radio is LPZ, and the frequency is 19.33 kHz [27].

In 1968, the Sanguine project, an extremely low-frequency long-wave radio station in the United States, was planned. However, the project needed to occupy nearly two-fifths of the area of Wisconsin, and it would be difficult for the huge antenna array planned by the project to survive under a nuclear attack. Finally, this project was not implemented [28].

The SHELF project was a deeply buried ELF communication system. It was closely followed by the SEAFARER project, which was a shallow-buried ELF antenna surface communication system. The two schemes were ultimately not implemented for various reasons [29]. The final implementation was Project ELF, which consisted of two sites located 148 miles apart in Wisconsin and Michigan, with an antenna array on the ground. The project was implemented in 1989, with a working frequency of 76 Hz and a standby frequency of 45 Hz. Because it was too large, vulnerable to fatal attacks, and local residents protested, the US Navy closed the only extremely low-frequency long-wave radio communication station in the United States in 2004 [28].

The frequency of the Russian ZEVS ELF long-wave radio is 82 Hz, and the output power of the entire ELF long-wave radio communication facility was approximately 10–14 MW [18]. India's ELF long-wave radio station was successfully built in 2012. At present, owing to military confidentiality and other reasons, little information is known about the call sign or frequency of this ELF long-wave radio station [30].

Unfortunately, it is difficult for underwater submarines to receive external signals. To communicate with submarines, countries spared no expenses in establishing long-wave radio stations. However, this communication mode had several limitations. The communication rate was extremely low. Therefore, under normal circumstances, long-wave communication was only used as a telephone to inform the submarine to surface quickly and then send out a communication buoy for rapid communication. However, such a situation is extremely easy to detect using radio positioning [31]. During World War II, Britain used radio positioning technology to carry out targeted attacks on German wolf-pack submarines [32].

Submarines and other underwater equipment typically use ultra-low and very low frequencies for limited communication, and the communication rate is only approximately 300 bit/s [33]. However, the generation of low-frequency, very-low-frequency, and extremely low-frequency signals requires extremely powerful

transmitting antennas, which are very expensive [34] and vulnerable to repeated damage in the communication process.

Because the conductivity of seawater is very good, it stimulates the induced current to consume the energy of electromagnetic waves, which seriously attenuates the RF signal during underwater transmissions. This limits the underwater communication distance of the RF signal; thus, it is only suitable for short-distance underwater communication, and cannot do the long-distance and high-speed underwater information transmissions needed in the future [35].

Optical Communication

The communication rates of underwater acoustic and radio frequency communication in water cannot reach Mbit/s and cannot transmit large-capacity high-definition videos. Owing to these limitations, optical communication has become a hot topic in underwater communication research because of its higher spectral efficiency.

Optical wireless communication has a high spectrum (400–800 THz) and is advantageous for high-speed communication. The communication distance can reach 100 m and the communication rate can reach the Gbit/s level [36]. Moreover, its volume is much smaller than the first two, its bandwidth utilization is high, its energy consumption is low, its beam directivity is good, its divergence angle is small, and its security is high, all of which are conducive to long-distance transmission [37]. Because the frequency of light waves is relatively high, their ability to carry information is strong, and a high-capacity wireless communication system can be built [38].

However, underwater optical communication has certain limitations. Because of the influence of underwater-channel scattering, absorption, turbulence, and other factors, the transmission distance is limited. Table 31.1 compares the advantages and disadvantages of different underwater wireless technologies.

This chapter briefly introduces the transmission characteristics of underwater optical wireless communication, summarizes and discusses the key technologies and progress in underwater optical wireless communication, including channel modeling, light source selection, coding modulation technology, and detection technology, and summarizes the difficulties and future development directions of underwater optical wireless communication.

Model and Research Status of Underwater Optical Wireless Communication Systems

System Model

An underwater optical wireless communication system includes transmitting equipment, an underwater channel, and receiving equipment. Figure 31.1 shows the underwater optical communication system model. The transmitting end is composed

Table 31.1 Advantages and disadvantages of different underwater wireless technologies

Underwater wireless technology	Advantages	Disadvantages
Acoustic	Relatively mature, with small attenuation and long transmission distance	Low communication rate, time delay, high-power consumption, poor confidentiality, and harmful to marine organisms
Radio frequency	Less affected by turbulence	Short communication distance, low communication frequency, severe attenuation, and high transmission power
Optical	High communication rate, small size, high-bandwidth utilization, and low-power consumption	Affected by absorption and scattering

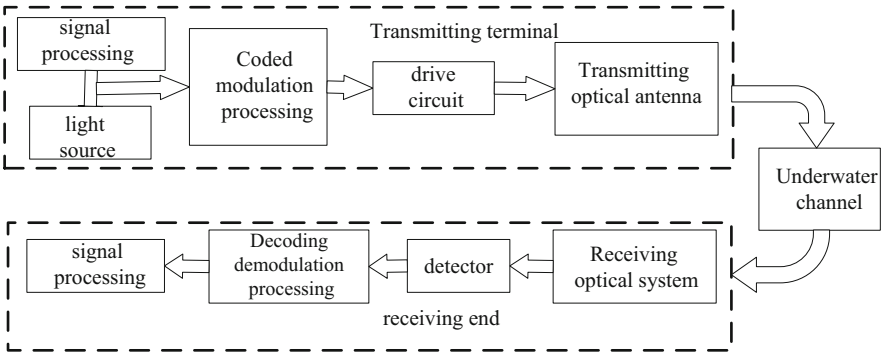


Fig. 31.1 Underwater optical wireless communication system

of a light source, coding–modulation processing circuit, driving circuit, and transmitting antenna. The receiver is composed of a detector, decoding–demodulation processing, filter, and output signal unit.

The underwater channel is relatively complex. In theory, an underwater channel includes turbulence, absorption and scattering factors, noise, and other factors affecting the communication.

The working process of the underwater optical communication system is as follows. At the transmitting end, the signal is encoded by an encoder, converted into a current that changes with the signal by a modulator, loaded onto the light source through a driving circuit, and finally transmitted through the transmitting antenna. At the receiving end, the receiving antenna gathers the received beam onto the detector surface. The detector converts the optical signal into an electrical signal, demodulates the signal, and finally recovers the original information.

Foreign Research Status

Since the 1960s, researchers have been exploring underwater optical communication systems [1]. In 1963, Duntley et al. [39] found that 470–570-nm blue-green light had the smallest attenuation in water, compared with other wavebands; that is, there was a low energy consumption "window." This discovery provided a reference for the selection of light sources for underwater optical communication systems.

In 1968, Gordon [40] used a laser as an emission light source underwater and experimentally proved the feasibility of using a laser to measure the depth of seawater. Moreover, he established the world's first system to measure the depth of seawater using a laser.

In 1972, Phillips et al. [41] introduced a technology to independently measure the absorption and scattering coefficients of seawater. This technology used an airborne system that transmitted laser pulses down to the water and recorded the upward scattering signals of particles in the water. The feasibility and effectiveness of the system were verified by comparing the experimental results with field measurement results.

In 1976, Karp [42] established optical wireless communication models for underwater and above-water terminals using multiple scattering models and experimentally proved their feasibility. The results showed that the radiation function was related to the transformation of the spatial covariance function (mutual interference function).

In 1977, Doherty et al. [43] used a focused blue-green laser diode as the light source to realize a one-way underwater optical wireless communication system from shore to submarine. This confirmed that it was feasible to establish an underwater optical wireless communication system, with a laser diode as the transmitter, to communicate at greater depths and higher cruising speeds.

In 1980, Wiener et al. [44] introduced a method for communicating underwater with a marine missile submarine. In the underwater submarine experiment, a blue-green LED was used as the light source of the transmitter to verify the feasibility of the system. This method was used in national defense and military affairs.

In 1983, the Soviet Union's Black Sea Fleet [45] realized underwater submarine communication near Sevastopol. The emission light source of this system was a blue-green laser. Finally, the system was used to implement a scheme for space-based and ground-based submarine communication.

In 1986, Billard et al. [46] studied an algorithm for airborne bathymeters. This algorithm could detect and locate reflections from the sea surface and seabed within 256 consecutive 2-ns received green light waveforms. The algorithm was flexible and simple and met the data processing requirements of an airborne bathymeter.

In 1992, Snow et al. [47] used an optical receiver to detect a green laser beam and measured the functional relationship between the propagation distance, data rate, and bit error rate. The maximum data rate was 50 kbit/s, and the error rate was 10^{-9} in a 50-m water tank.

In 1995, Bales et al. [48] designed an underwater optical communication system that used an LED as the emission light source. The experiment proved that the

underwater optical communication system had a communication rate of 1–10 kbit/s and a communication distance of 20–30 m in clear water.

In 2000, Tivey et al. [49] developed underwater short-distance communication equipment for submarine communication with a blue LED as the emission light source and a silicon photodiode as the detector at the receiving end. Error-free reception was realized in a 2.7-m water tank, and errors occurred when the communication distance exceeded 3 m. Finally, an underwater experiment with a communication distance of 5 m, communication rate of 115.2 kbit/s, and 70% accuracy rate was realized in the pool.

In 2003, Schill et al. [50] established a small, real-time optical communication transceiver using a group of small autonomous submersibles. The transceiver used a 3-W high-power blue-green laser diode and a 520-nm photodiode. An underwater experiment with a 2-m communication distance, 57.6-kbit/s data transmission rate, and no error codes was successfully implemented.

In 2004, Akyildi et al. [51] used a high-power blue-green LED as a light source in an underwater optical communication system. The experimental results showed that the system could complete underwater experiments with a 1.7-m communication distance, 57.6-kbit/s data transmission rate, and almost no error codes.

In 2005, Farr et al. [52] proposed a preliminary design for an optical modem system based on omnidirectional sources and receivers. First, the functional requirements and system constraints of the system driver were reviewed. The optical transmission characteristics of seawater and the factors affecting high-speed communication were discussed. The relationship between the link budget, data rate, and distance was established. In the experiment, a blue light-emitting diode was used as the light source, the detector was a photomultiplier tube, and the switch keying modulation mode was used. A vertical ocean test experiment of nearly 100 m was conducted, and the communication rate was close to 10 Mbit/s.

In 2007, Cochenour et al. [53] introduced the influence of scattering on underwater light transmissions. Using binary phase shift keying, quadrature phase shift keying, 8-phase shift keying, 16 quadrature amplitude modulation, and 32 quadrature amplitude modulation, an underwater experiment were conducted in a turbid water environment. The data rate was up to 5 Mbit/s, communication distance was 3.6 m, and bit error rate was 10^{-3}.

In 2008, Hanson et al. [54] simulated an underwater signal transmission system with a 532-nm laser diode in a 2-m water pipe in a laboratory. The 532-nm light source originated from a 1064-nm continuous-wave (CW) laser diode. Using intensity modulation and Monte Carlo simulations, the extinction ratio was 36 dB. An underwater experiment with a communication rate of 1 Gbit/s, communication distance of 2 m, and almost no error codes was conducted.

In 2008, Cox et al. [55] used a 405-nm laser diode as the emission light source and used switch keying modulation and a Reed–Solomon (RS) (255, 129) coding mode to successfully conduct underwater transmission experiments with a communication rate of 500 kbit/s, communication distance of 3.66 m, and a bit error rate lower than 10^{-4}. The experimental results showed that the signal-to-noise ratio of the coded system was improved by 8 dB, compared to that of the uncoded system.

In 2008, Jaruwatanadilok [56] used the vector radiative transfer theory to model an underwater optical communication system. A 532-nm laser diode was used as the emission light source, and the Monte Carlo simulation method was used to successfully realize an underwater communication experiment. It had a communication rate of 1 Gbit/s over 30 m, a communication rate of 500 Mbit/s over 50 m, and no error codes in the outbound communication.

In 2009, Doniec et al. [57] used an AquaOptical device for experiments with an underwater communication system. In clear water, the communication rate was 1.2 Mbit/s over a 30-m transmission distance, with a received data success rate of 100%. An underwater communication experiment with a 0.6-Mbit/s data rate and 9-m communication distance was realized in turbid water with a visibility of 3 m; however, the accuracy of the received data decreased.

In 2009, Cochenour et al. [58] studied an underwater communication system using a beam spread function. The system used a 520-nm light-emitting diode at the transmitter and a photomultiplier tube as the detector at the receiver. The Monte Carlo simulation method was used, and a small-angle approximation method was used to solve the radiative transfer equation. An underwater communication experiment with a transmission rate of 70 Mbit/s, communication distance of 107 m, and almost no error codes was realized.

In 2010, Simpson et al. [59] adopted digital signal processing and error correction technology. In the experiment, a 470-nm LED was used as the emission light source, and a photomultiplier tube was used as the detector. They were tested in 3-m and 7.7-m water tanks. The test speed was 5 Mbit/s. Under a 10^{-5} bit error rate, the signal-to-noise ratio was improved by 4 dB, compared with the uncoded data.

In 2010, Cox et al. [60] used a 532-nm blue-green laser diode as the emission light source and adopted orthogonal frequency shift keying modulation and an RS (255, 129) coding mode. They conducted a MATLAB simulation experiment in a 7.7-m underwater channel. When the transmission rate was 500 kbit/s, there were almost no errors. When the transmission rate was 1 Mbit/s, the error rate was 1.19×10^{-2}.

In 2011, Vucic et al. [61] used a high-power white LED as the emission light source and an avalanche photodiode with quadrature amplitude modulation as the detector in underwater experiments. The data transmission rate was 513 Mbit/s and the bit error rate of the received data was 2×10^{-3}.

In 2012, Gabriel et al. [62] used a 532-nm LED transmitter and a silicon PIN photodiode receiver, using a switch keying modulation method, and used the bit error rate as the standard for evaluating the system. An underwater experiment with a data transmission rate of 1 Gbit/s, communication distance of 20 m, and 10^{-6} bit error rate (BER) was realized.

In 2013, Cossu et al. [63] used two low-cost LED arrays as the emission light source and avalanche photodiodes as the detector. Using the Manchester coding method, they successfully conducted an underwater experiment with a 6.25 Mbit/s communication rate in a 3.3-m-diameter outdoor water tank. When the discrete multitone modulation mode was adopted in the experiment, a communication rate

of up to 58 Mbit/s was realized. The experimental results showed that error-free transmission could be realized under all experimental conditions.

In 2013, Doniec et al. [64] designed an end-to-end underwater communication system. The system used a 470-nm LED as the emission light source, an avalanche photodiode as the detector at the receiving end, and switch keying modulation and RS (255, 223) coding. They successfully conducted underwater communication experiments with a 10 Mbit/s communication rate, 36-m communication distance, and 10^{-6} error rate.

In 2014, Han et al. [65] designed an acousto-optic hybrid underwater communication system and estimated its performance through the QualNet simulator. The system used an AquaOptical II optical modem, which could communicate underwater with a maximum communication distance of 50 m, maximum communication rate of 2.28 Mbit/s, and almost no error codes. The system also used a S2CR18/34 acoustic modem, which realized underwater communication experiments with a maximum communication distance of 3.5 km, maximum communication rate of 13.9 kbit/s, and almost no error codes.

In 2014, Hiskett et al. [66] designed an underwater communication system with a 450-nm laser diode as the light source at the transmitter and a silicon avalanche photodiode at the receiver. The experimental results showed that the data transmission rate was up to 45.32 Mbit/s and the bit error rate was 1.8×10^{-4} in a 1-m-long water tank.

In 2015, Oubei et al. [67] adopted a TO-9 packaged 450-nm braided laser diode and the 16 orthogonal-amplitude modulation orthogonal frequency-division multiplexing (16 OAM-OFDM) modulation mode to achieve a transmission distance of 5.4 m, communication rate of 4.8 Gbit/s, bit error rate of 2.6×10^{-3}, and signal-to-noise ratio of 15.63 dB.

In 2015, Nakamura et al. [68] used the intensity modulation/direct detection orthogonal frequency-division multiplexing method with avalanche photodiodes as detectors. An experiment with a 405-nm blue laser diode transmission in a 4.8-m underwater channel was completed, with a communication rate of up to 1.45 Gbit/s and a bit error rate of 9.4×10^{-4}.

In 2016, Oubei et al. [69] used commercially available TO-9 packaged 520-nm braided laser diodes and avalanche photodiodes with switch keying modulation in underwater experiments. The transmission distance was more than 7 m, the communication rate was up to 2.3 Gbit/s, and the bit error rate was 2.23×10^{-4}.

In 2016, Stephen et al. [70] adopted a 420-nm laser diode doped with AlGaInN, and the maximum optical power of the laser diode was 100 mW. In the laboratory water tank, an underwater experiment was conducted with a data transmission rate of up to 2.5 Gbit/s, 10-m transmission distance, and almost no error codes. GaN laser diodes have been proven to be suitable for underwater communication systems.

In 2017, Abdullah et al. [71] used commercially available TO-9 packaged 520-nm laser diodes as emission light sources and avalanche photodiodes as detectors. An underwater video-image transmission experiment was successfully conducted with a communication rate of 4 Gbit/s and a 5-m communication distance in clear and port seawater. When the modulation modes of quadrature phase shift

keying (QPSK), quadrature amplitude modulation (QAM), and 8-quadrature amplitude modulation (8-QAM) were adopted, the bit error rate was 10^{-9}. When the 8-PSK modulation mode was adopted, the bit error rate was 9.9×10^{-9}.

In 2017, Tsai et al. [72] used a 450-nm blue GaN laser diode as the emission light source and a PIN photodiode as the detector and adopted the 16 OAM-OFDM mode in an underwater communication experiment. The communication rates were 12.4, 9.2, and 5.2 Gbit/s, respectively; the communication distances were 1.7, 6.8, and 10.2 m, respectively, and the bit error rates were 2.7×10^{-3}, 2.9×10^{-3}, and 3.2×10^{-3}, respectively.

In 2017, Oubei et al. [73] modeled an underwater channel with a temperature gradient and adopted an experimental method with a generalized gamma distribution. The experimental results showed that the measured data for all underwater channel conditions had a good fit. In addition, it was observed that a thermally homogeneous channel can be well described by a simple gamma distribution. This was the first model to comprehensively describe thermal homogeneous and gradient-based underwater optical channels.

In 2017, Jamali et al. [74] used multiple-input multiple-output (MIMO) spatial diversity technology on an underwater optical wireless communication link to mitigate the fading caused by turbulence. In the experiment, a 450-nm LED was used as the emission light source, the switch keying modulation mode was adopted, and the Monte Carlo simulation method was used to realize an underwater experiment in a 25-m pool in the laboratory. The bit error rate was lower than 10^{-9}. Finally, considering the effects of absorption and scattering, the exact expressions of the optimal gain, equal gain combination, and upper bound of the BER were obtained using an analytical solution method.

In 2018, Halafi et al. [75] designed an underwater video transmission system that could realize two-way underwater communication. The uplink and downlink emission light sources of the system were a commercial TO-9 packaged 450-nm blue laser diode and 520-nm green laser diode, respectively, with a bandwidth of 1.2 GHz. The uplink and downlink receiver detectors used avalanche photodiodes. When the communication distance was 4.5 m, the communication rate reached 15 Mbit/s when using the quadrature phase shift keying (QPSK) modulation mode. When 16-QAM and 64-QAM were adopted, the communication rate reached 30 Mbit/s, and the peak signal-to-noise ratio in the harbor seawater reached 16 dB.

In 2018, Illi et al. [76] used an improved radiative transfer equation to predict the optical path loss in an underwater channel and calculated the change in the BER performance of an underwater optical wireless communication system with the transmission time and distance. In the experiment, an LED was used as the emission light source, and the intensity modulation/direct detection (IM/DD) method was adopted. An underwater experiment with a communication distance of 3.5 m and a BER of less than 10^{-4} was successfully realized.

In 2019, Illi et al. [77] further improved the method for calculating the radiative transfer equation in an underwater communication system. The system used an LED as the emission light source and realized an underwater experiment with a communication distance of 6 m and a BER of 10^{-6}. The experimental results showed that

the improved method of the radiative transfer equation obtained the same results by calculating the BER of an underwater optical wireless communication system. It was consistent with the Monte Carlo simulation results, thereby reducing the system complexity.

In 2019, Tsai et al. [78] proposed and demonstrated an underwater wireless laser transmission system using four-level pulse amplitude modulation (4-PAM). The system was equipped with a beam constrictor/expander on a 12.5-m pipeline underwater channel and 2.5-m high-turbidity port underwater channel, respectively, in an underwater experiment. It achieved data transmission rates of 15 Gbit/s and 30 Gbit/s and BERs of 2.4×10^{-7} and 10^{-6}, respectively. In the highly turbid underwater link of the harbor, the large amount of scattered light received by the optical receiver of the 4-PAM system improved with the beam-shrinking/expanding device.

In 2020, Ramavath et al. [79] used the switch keying modulation method and the hyperbolic tangent distribution method to specify the probability density function of a lognormal distribution. The new closed-form analytical BER expressions for single-input single-output (SISO), single-input multiple-output (SIMO), multiple-input single-output (MISO), and MIMO links in the noncoding and RS coding cases were calculated and verified using Monte Carlo simulation results. The underwater data transmission rate was 500 Mbit/s, the communication distance was 30 m, and the BER of the received data was 10^{-5}. It realized underwater Internet of things applications, such as communication between submarines, between submarines and docks, between sensors collecting ocean parameters, and aggregation equipment.

In 2020, Arvanitakis et al. [80] used six 450-nm LED arrays as emission light sources, PIN photodiodes as detectors at the receiving end, and the switch keying modulation mode in an underwater communication experiment. They successfully achieved communication rates of 4.92, 3.22, and 3.4 Gbit/s; communication distances of 1.5, 3, and 4.5 m; and bit error rates of 1.5×10^{-3}, 1.1×10^{-3}, and 3.1×10^{-3}, respectively.

In 2022, Kong et al. [81] proposed a prototype for a real-time digital video dynamic underwater monitoring system. The system used a 458-nm LED as the light source, and the video stream was processed by a transmitter circuit composed of video coding and switch keying modulation. In a 1.5-m pure water channel, 1.53 mL/s bubble-induced turbulent pure water channel, and other channels, the communication rate was 2.5 Mbit/s and the bit error rate of the received data was 5.31×10^{-4}.

The foreign research status of underwater optical communication system models is summarized in Table 31.2.

The foreign research progress on underwater optical communication rates (kbit/s, Mbit/s, and Gbit/s) is summarized in Table 31.3.

It can be seen from the foreign research status of underwater optical wireless communication systems that, to improve the communication performance, researchers have attempted to improve the communication rates and increase the transmission distances through optimizing the modulation modes and light sources of underwater optical communication systems. They have also analyzed underwater optical wireless communication systems through various simulation models,

Table 31.2 Foreign research progress on underwater optical communication system models

Year	Researcher	Research content	Research results
1963	Duntley	Attenuation of light of different wavebands in water	Blue-green light with a 470–570 nm wavelength has the weakest attenuation in water
1968	Gordon et al.	Characteristics of laser propagation under water	Established the world's first system for measuring the depth of seawater using a laser
1972	Phillips et al.	Technique for independently measuring the absorption and scattering coefficients of seawater	The experimental results were compared with the field measurement results to prove the feasibility and effectiveness of the system
1976	Karp et al.	Established an underwater optical wireless communication model above the water surface	The feasibility of the system was proved using multiple scattering models
1977	Doherty et al.	Established an underwater communication system using blue-green laser diodes	Realized a one-way underwater optical wireless communication system from shore to submarine
1980	Wiener	Research on the communication modes of marine missile submarines	The feasibility was verified by using blue-green light communication in an underwater submarine experiment
1983	Soviet Black Sea Fleet	Research on submarine communication using blue-green lasers	Scheme design of space-based to submarine communication and ground-based to submarine communication was carried out
1986	Billard et al.	Algorithm was studied for the WRELADS airborne bathymeter	The bathymeter could detect and locate the reflections of the sea surface and sea bottom in 256 consecutive 2-ns received green light waveforms

compared the simulation data of communication rates and distances with engineering tests, and verified the derivation of theoretical data through experiments.

Domestic Research Status

China began to study underwater optical communication technology for use against submarines in the 1980s. In 1987, Jiang et al. [82] designed a special codec image compression scheme. The experimental results showed that the image transmission rate reached 640 kbit/s and the bit error rate was 5.2×10^{-2} after using the codec image compression scheme. A complete image could be transmitted in about 10 s.

In 1989, Ao [83] used pulse position modulation in an experiment and RS coding in a laser submarine communication channel. He analyzed the error correction ability of RS (7, 3), (15, 7), (31, 15), (63, 31), and (255, 223) codes in the channel. The results showed that the RS error correction coding method could be applied to a laser

Table 31.3 Foreign research progress on communication rates (kbit/s, Mbit/s, and Gbit/s)

Classification	Year	Communication rate	Researcher	Communication distance	Error rate
Kbit/s	1992	50 kbit/s	Snow et al.	50 m	10^{-9}
	1995	1/10 kbit/s	Bales et al.	30/20 m	–
	2000	115 kbit/s	Tivey et al.	5 m	–
	2003	57.6 kbit/s	Schill et al.	2 m	Accuracy: 70%
	2004	57.6 kbit/s	Akyildi et al.	1.7 m	No error codes
	2008	500 kbit/s	Cox et al.	3.66 m	$< 10^{-4}$
	2010	500 kbit/s	Cox et al.	7.7 m	Almost no error codes
	2014	13.9 kbit/s	Han et al.	3.5 m	Almost no error codes
Mbit/s	2005	10 Mbit/s	Farr et al.	100 m	Almost no error codes
	2007	5 Mbit/s	Cochenour et al.	3.6 m	10^{-3}
	2008	500 Mbit/s	Jaruwatanadilok	50 m	No error codes
	2009	1.2/0.6 Mbit/s	Doniec et al.	30/9 m	Almost no error codes
	2009	70 Mbit/s	Cochenour et al.	107 m	Almost no error codes
	2010	1 Mbit/s	Cox et al.	7.7 m	1.19×10^{-2}
	2010	5 Mbit/s	Simpson et al.	7.7 m	10^{-5}
	2011	513 Mbit/s	Vucic et al.	–	$< 2 \times 10^{-3}$
	2013	58 Mbit/s	Cossu et al.	3.3 m	Almost no error codes

	2013	10 Mbit/s	Doniec et al.	36 m	10^{-6}
	2014	45.32 Mbit/s	Hiskett et al.	1 m	1.8×10^{-4}
	2014	2.28 Mbit/s	Han et al.	50 m	Almost no error codes
	2018	30 Mbit/s	Halafi et al.	4.5 m	–
	2020	500 Mbit/s	Ramavath et al.	30 m	10^{-5}
	2022	2.5 Mbit/s	Kong et al.	1.5 m	5.31×10^{-4}
Gbit/s	2008	1 Gbit/s	Hanson et al.	2 m	Almost no error codes
	2008	1 Gbit/s	Jaruwatanadilok	50 m	No error codes
	2012	1 Gbit/s	Gabriel et al.	20 m	10^{-6}
	2015	1.45 Gbit/s	Nakamura et al.	4.8 m	9.4×10^{-4}
	2016	2.3 Gbit/s	Oubei et al.	7 m	2.23×10^{-4}
	2016	2.5 Gbit/s	Stephen et al.	10 m	Almost no error codes
	2017	4 Gbit/s	Abdullah et al.	5 m	10^{-9}
	2017	12.4/9.2/5.2 Gbit/s	Tsai et al.	1.7/6.8/10.2 m	$2.7 \times 10^{-3}/2.9 \times 10^{-3}/3.2 \times 10^{-3}$
	2017	4.8	Oubei et al.	5.4 m	2.6×10^{-3}
	2019	15/30 Gbit/s	Tsai et al.	12.5/2.5 m	$2.4 \times 10^{-7}/10^{-6}$
	2020	4.92/3.22/3.4 Gbit/s	Arvanitakis et al.	1.5/3/4.5 m	$1.5 \times 10^{-3}/1.1 \times 10^{-3}/3.1 \times 10^{-3}$

submarine communication system to correct errors and greatly improve the channel error rate.

In 1992, Sui et al. [84] established a method for analyzing underwater optical communication systems and conducted experimental simulations using different modulation methods. With a signal-to-noise ratio of 30 dB, when using the switch keying modulation mode, the bit error rate was 10^{-2}. When the four-pulse position modulation (4-PPM) mode was used, the bit error rate was 10^{-3}. When the 8-PPM mode was used, the bit error rate was 10^{-5}. When the relative phase shift keying (RPSK) modulation mode was used, the bit error rate was 10^{-8}. However, when the SNR was less than 10 dB, the bit error rates of the several modulation modes were essentially the same.

In 1995, Wang et al. [85] studied a time slot synchronizer. An analysis of its performance showed that the synchronizer had a fast signal acquisition speed, strong noise resistance, etc., and could improve the performance of underwater communication systems.

In 1997, Yang et al. [86] introduced an underwater laser communication system using error correction code methods. The detector at the receiving end of the system used a photomultiplier tube, RS coding, and PPM to compare different error correction code methods based on an error rate analysis. The experimental results showed that when the input bit error rate was 2×10^{-2}, using the RS (255, 223) error correction code reduced the bit error rate, after decoding, to less than 10^{-6}, and the coding gain exceeded 1 dB.

In 1998, Huang et al. [87] studied a full-duplex underwater laser communication experimental device controlled by an 89C51 single-chip microcomputer in a 3-m-long water tank. The experiment showed that the maximum data transmission rate of the system could reach 21.2 kbit/s, and the received data bit error rate was 2.8×10^{-3}.

In 1999, Ao et al. [88] developed an underwater laser communication model using an adaptive array signal intensifier and diversity reception. The experimental results showed that the output signal-to-noise ratio of the adaptive array signal intensifier was proportional to the number of array input signals, which meant that the underwater adaptive array signal was less affected by the noise in the channel.

In 2002, He et al. [89] designed an underwater communication system using PPM with a 532-nm light wavelength, multiple diversity reception at the receiving end, and a PIN photodiode detector. The communication distance of the system was 50 m, and the bit error rate was 1.76×10^{-3} in a water tank in the laboratory. Underwater experiments were conducted in three different sea areas in Hainan Province, China, which further verified the feasibility of the system.

In 2003, Wang et al. [90] studied an underwater communication system using fiber-grating filters. The fiber wavelength used in the experiment was 532 nm; it achieved 100% reflectivity and a reflection bandwidth lower than 0.01 nm. This method expanded the application field of fiber gratings and was a new research method in underwater optical communication.

In 2004, Pang [91] successfully conducted an experiment with a 1 Mbit/s communication rate in a 10-m water tank in the laboratory using a blue LED. An

underwater experiment with a 13-m communication distance and 3 Mbit/s communication rate was successfully realized using an ultra-bright blue LED. The experimental results showed that the received file had an error rate of 8×10^{-9} with almost no bit errors.

In 2006, Ao et al. [92] designed a highly sensitive digital optical receiver for an underwater communication system based on digital PPM and digital signal processing technology. Experiments and tests showed that the receiver reduced the system's signal-to-noise ratio requirements and achieved good performance.

In 2008, Cui [93] used a 532-nm green laser diode as the transmitting light source in an underwater communication system. Polarized light was used as the carrier, and the frequency of the signal pulse was 128 Hz. In the laboratory water tank, he performed an experiment with a 5-m communication distance, which proved that using polarized light as the carrier could reduce the bit error rate and improve the communication rate of the underwater optical communication system.

In 2008, Luo et al. [94] used a blue-green LED as the emission light source and designed an underwater optical communication system based on the Infrared Data Association (IrDA) physical-layer communication protocol. The system used a curve-fitting method to establish the relationship between the light intensity and distance. The experimental results showed that an underwater communication experiment with a data transmission rate of 125 kbit/s and error-free code could be realized without any influence of the light source within 2 m. This verified that the system could meet the high-speed communication requirements of underwater mobile wireless sensor network nodes.

In 2009, He et al. [95] designed a heterodyne laser Doppler vibrometer with an optical fiber structure to analyze the impact of water waves on an acousto-optic communication system. The experiment showed that the acousto-optic communication system could achieve a communication rate of 5 kbit/s without waves on the water surface, and there were almost no error codes in the experiment. With small waves on the water surface, the system could complete the experiment with a communication rate of 30 kbit/s and almost no error codes.

In 2010, Liu et al. [96] simulated the performance of the laser uplink of an underwater platform received by a satellite using the Monte Carlo simulation method. Using the PPM and maximum-likelihood estimation methods, they calculated that the bit error rate of the communication system was less than 10^{-4} and proposed a laser uplink communication scheme between 60-m underwater and a satellite.

In 2011, Cong [97] introduced the structure of an adaptive underwater optical communication network framework and studied an adaptive decoding scheme applied to underwater optical communication systems based on different coding schemes. The data transmission rate was 3 Mbit/s, the communication distance was 20 m, and the bit error rate was 4×10^{-4} in the underwater experiments.

In 2012, Wang et al. [98] studied an underwater optical communication system that could compress the dynamic range of the signal and used a variable gain amplifier to vary the dynamic range of the compressed signal, so that the signal eventually stabilized. This method not only reduced the dynamic range of the signal

but also effectively improved the accuracy of the digital-to-analog converter, which could be well applied to underwater optical communication systems.

In 2012, Ao et al. [99] designed a control system that used an electronic compass to adjust the angle of an underwater optical transmitter. Through the electronic compass, the depression and elevation angles of the system at the current time were obtained so that the transmitter could be adjusted to an appropriate angle, according to the actual situation. The system realized a transmission experiment with a communication distance of approximately 10 m and no error codes.

In 2013, Tian et al. [100] designed and implemented an underwater optical communication system based on LEDs. In the experiment, the RS232 serial communication protocol was used to encode and decode information. By adjusting the communication baud rates at different speeds, the final experimental results showed that the data transmission rate was 38.4 kbit/s and the communication distance was 30 m.

In 2013, Xu et al. [101] introduced an underwater optical communication system with IM/DD orthogonal frequency-division multiplexing (OFDM). The system used a blue light-emitting diode as the emission light source and a PIN photodiode as the receiving detector. In the 2-m underwater channel, 16-QAM, 32-QAM, and 64 QAM modes were used to achieve communication rates of 161.36 Mbit/s, 156.31 Mbit/s, and 131.07 Mbit/s, respectively. The bit error rates were, respectively, 2.5×10^{-3}, 7.42×10^{-4}, and 3.17×10^{-3} in underwater experiments.

In 2013, Hu et al. [102] designed a high-speed underwater communication system using digital signal processing. The light source at the transmitting end of the system used a 532-nm laser diode, and the detector at the receiving end used a photomultiplier tube. Using PPM and RS coding, an error-free full-duplex underwater communication experiment with a 73-kbit/s communication rate was realized in a 100-m pool.

In 2014, Xi et al. [103] produced a frequency response model of an underwater optical wireless communication channel using the discrete coordinate method based on the vector radiative transmission theory. They used this model to evaluate the performance of an underwater optical communication system using m-ary PPM. The experimental results showed that the communication rate was 1 Gbit/s, and the bit error rate performance decreased with an increase in transmission distance. The BER performance of quaternary pulse position modulation was almost the same as that of switch keying modulation. Octal pulse position modulation or 16 PPM could be used to further improve the bit error rate performance.

In 2015, Cai et al. [104] used blue-green light-emitting diodes as the emission light source and photodiodes as the receiver detector in an underwater communication system. In a 10-m water tank in the laboratory, they conducted an experiment with a data transmission rate of 115.2 kbit/s and almost no bit errors.

In 2016, Shen et al. [105] realized a small, low-power underwater communication system using a single-mode 450-nm laser diode, a silicon avalanche photodetector, and a simple non-return-to-zero switch keying modulation mode. In a 12-m underwater channel, the received data had a bit error rate of 2.8×10^{-5} and a

communication rate of up to 2 Gbit/s. In a 20-m underwater channel, the received data had a bit error rate of 3.0×10^{-3} and a communication rate of up to 1.5 Gbit/s.

In 2016, Hu et al. [106] studied an underwater two-way communication system using a blue LED. The system used a photodiode as the receiver detector and realized an error-free full-duplex underwater communication experiment with a communication distance of 5 m and communication frequency of up to 60 MHz.

In 2016, Zhou et al. [107] used a blue-green laser diode as the emission light source of an underwater system. The detector used an avalanche photodiode to combine low-density parity check codes and PPM to form a simplified soft demodulation method. MATLAB simulation verified that the bit error rate performance of the simplified soft demodulation method was better than that of the RS code, proving the effectiveness of the method in an underwater optical communication system.

In 2016, Lu et al. [108] used a 405-nm blue laser diode as the emission light source, an avalanche photodiode as the detector at the receiving end, and adopted the 16 OAM-OFDM mode to successfully conduct an underwater experiment with a data transmission rate of 9.6 Gbit/s, communication distance of 8 m, and bit error rate of 10^{-4}.

In 2016, Xu et al. [109] studied the transmission characteristics of red light in water from the perspective of the extinction coefficient and channel bandwidth using the Monte Carlo simulation method. In the laboratory, a simple low-cost top 5 red laser diode was used as the emission light source and an avalanche photodiode was used as the detector. They used 128- and 32-OAM-OFDM modes in underwater experiments. In a 6-m-long water tank, the data transmission rates were 1.324 Gbit/s and 4.883 Gbit/s, respectively, and the bit error rates were 2.02×10^{-3} and 3.20×10^{-3}, respectively.

In 2017, Liu et al. [110] used a 520-nm green laser diode and the switch keying modulation mode in an underwater experiment. The communication distance was 34.5 m, the communication rate was up to 2.7 Gbit/s, and the received data bit error rate was less than 3.8×10^{-3}. In the experiment, the maximum data rate reached 4.60, 4.20, 3.93, 3.88, and 3.48 Gbit/s, when the underwater distances were 2.3, 6.9, 11.5, 19.1, and 31.7 m, respectively.

In 2017, Chen et al. [111] used a single-mode braided green laser diode as the light source, and the modulation technology employed 32 OAM-OFDM. They successfully completed a 5.3 Gbit/s uplink transmission with no power load in a 21-m underwater channel and the received data bit error rate was 2.47×10^{-3}. They also completed a 5.5 Gbit/s downlink transmission with a power load, and the received data bit error rate was 2.92×10^{-3}.

In 2017, Ao et al. [112] used a blue-green LED as the emission light source with pulse position modulation. The receiver detector used photodiodes. They successfully completed an underwater experiment with multiple inputs and outputs with a data transmission rate of 10 Mbit/s, a communication distance of 2.3 m, and error-free codes.

In 2017, Lin [113] used a green laser diode as the emission light source and a PIN photodiode as the detector at the receiving end, adopted OFDM, and used a field

programmable gate array. They achieved a communication distance of 2 m, a communication rate of 80 Mbit/s, and a bit error rate of 2.24×10^{-3}.

In 2017, Song Aerospace [114] used a blue light-emitting diode as the emission light source and a photodiode as the detector at the receiving end and adopted the MIMO OFDM mode. They realized underwater communication experiments using MATLAB programming with a 2-m communication distance, a communication rate of 31.082 Mbit/s, a signal-to-noise ratio of 20 dB, and a bit error rate lower than 10^{-4}.

In 2018, Huang et al. [115] adopted the 16-OAM frequency-division multiplexing modulation mode and used a blue laser diode as the emission light source. They achieved a communication distance of 1.7 m, and a received data bit error rate of 1.9×10^{-5}; the data transmission rate was 14.8 Gbit/s. When the communication distance was 10.2 m, the bit error rate of the received data was 1.5×10^{-4}, and the data transmission rate was 10.8 Gbit/s.

In 2018, Wang et al. [116] designed an underwater optical communication system using PPM. The system used a 450-nm blue light-emitting diode as the emission light source and an avalanche photodiode as the detector at the receiving end. An undistorted transmission experiment with a 6-m communication distance and 0.98 Mbit/s communication rate was realized in the laboratory.

In 2018, Han [117] used a 472-nm blue LED as the emission light source, a photomultiplier tube as the receiver detector, switch keying modulation, and RS coding. Using the Monte Carlo simulation method, he successfully realized an underwater two-way network video transmission experiment with the longest communication distance of 9.6 m, the largest communication rate of 20 Mbit/s, and a bit error rate of 10^{-9}.

In 2018, Yang et al. [118] used a high-power blue-green LED as the emission light source, a photomultiplier tube as the detector, and PPM in an underwater communication system. In the laboratory water tank, they successfully completed a 3-m communication distance, a data transmission rate of 6 Mbit/s, and a bit error rate of 7×10^{-4}.

In 2018, Wang et al. [119] used a low-cost 448-nm blue LED as the emission light source and an avalanche photodiode as the detector at the receiving end. An underwater communication experiment with a 25 Mbit/s communication rate, a 10 m communication distance, and a bit error rate of less than 10^{-4} was successfully realized using a non-return-to-zero modulation mode on the key control and an RS (255, 223) coding mode.

In 2018, Zhang [120] used a 520-nm laser diode as the emission light source, a photodiode as the detector at the receiving end, and PPM to successfully achieve underwater video and audio transmission experiments in the water tank of the laboratory. They achieved a communication rate of 9.6 kbit/s, a communication distance of 3 m, and an error rate of less than 10^{-5}.

In 2018, Li [121] used a 515-nm green LED as the emission light source and an avalanche photodiode as the detector at the receiving end. They designed an underwater optical wireless two-way communication system using an Ethernet conversion interface module. The system achieved a communication rate of

100 Mbit/s, communication distance of 16 m, and bit error rate of 2.4×10^{-6}, when transmitting voice, files, real-time video, and other services.

In 2019, Wang et al. [122] used a 520-nm green laser diode in an underwater communication system using switch keying modulation. In experiments using tap water, they successfully achieved a communication rate of up to 500 Mbit/s, a communication distance of 100 m, and a bit error rate as low as 2.5×10^{-3}. According to the measured minimum required power and power attenuation model in the underwater channel, the predicted transmission rates were 500 Mbit/s and 100 Mbit/s, and the communication distances were 146 m and 174 m, respectively.

In 2019, Hong et al. [123] used a 450-nm laser diode, discrete multitone modulation, and probabilistic constellation shaping (PCS) technology to approximate the channel capacity limit of a discrete multitone transmission. The available modulation bandwidth was 2.75 GHz. Underwater experiments with communication distances of 5 m, 25 m, and 35 m and net data transmission rates of 18.09 Gbit/s, 19.21 Gbit/s, and 12.62 Gbit/s, respectively, were realized. Compared with conventional QAM, the capacity increased by 32.22%, 30.03%, and 31.55%, respectively.

In 2019, Jiang et al. [124] carried out detection experiments in the South China Sea using dual-wavelength multi-channel lidar consisting of 486.1-nm and 532-nm blue-green bands. The experimental results showed that the laser radar could achieve a maximum communication distance of 160 m, and the communication rate was 30% higher than that of the traditional 532-nm blue-green laser diode.

In 2019, Wang et al. [125] used a 450-nm laser diode as the emission light source, a PIN photodiode as the detector at the receiving end, adopted 16 OAM and used a field programmable gate array to conduct an underwater high-speed communication experiment. They achieved a communication rate of 50 Mbit/s, a communication distance of 3 m, and a bit error rate of 7.11×10^{-4}.

In 2019, Zhou et al. [126] used a GaN-doped LED as the emission light source and PIN photodiode as the detector at the receiving end. A 64-QAM discrete multitone modulation mode was adopted, and an underwater communication experiment with a communication rate of 14.81 Gbit/s was successfully implemented in a 1.2-m-long water pipe in the laboratory. With bit-loaded discrete multitone modulation, an underwater communication experiment with a communication rate of 15.17 Gbit/s was also realized. The bit error rate under both modulation modes was lower than 3.8×10^{-3}.

In 2019, Meng [127] used a 532-nm laser diode as the emission light source, a PIN photodiode as the receiver detector, and the switch keying modulation mode. The longest underwater transmission distance was 12 m and the bit error rate was 9.028×10^{-4} for text transmissions and 5.031×10^{-4} for 640 × 480 resolution video and picture transmissions.

In 2019, Han [128] used a 450-nm laser diode as the emission light source and a photomultiplier tube as the detector at the receiving end. Using switch keying modulation and nonlinear compensation, he successfully realized an underwater high-speed communication experiment with a communication rate of 6.23 Gbit/s, a communication distance of 66 m, and a bit error rate lower than 10^{-4}.

In 2019, Shang [129] used a blue-green laser diode as the emission light source and a photodetector at the receiving end. Using switch keying modulation and RS coding, he built an underwater real-time optical wireless communication experimental system. The experimental results showed that the system had a communication rate of 75 Mbit/s, communication distance of 4 m, and no error codes.

In 2020, Li et al. [130] designed a communication system that could realize two-way underwater video transmissions. The system used a 445-nm light-emitting diode as the light source and an avalanche photodiode as the detector at the receiving end for a full-duplex underwater video transmission experiment in the water tank of the laboratory. It adopted binary frequency shift keying and RS (255, 239) coding. The communication rate was 1 Mbit/s, the communication distance was 10 m, and the bit error rate was 3.9×10^{-6}.

In 2020, Ji et al. [131] proposed a low-complexity two-level chaotic encryption scheme that could be used in underwater optical wireless communication and extended discrete multitone modulation technology using a discrete Fourier transform. The experimental results showed that encryption did not negatively impact the performance of the proposed underwater optical communication system. An underwater experiment with communication distances of 50 m and 55 m and communication rates of 5 and 4.5 Gbit/s, respectively, was successfully implemented.

In 2020, Ao et al. [132] introduced a method to restore distorted underwater images using a depth learning algorithm. The simulation results showed that this method could effectively restore images transmitted underwater. This algorithm increased the transmission rate of the image. The average running time was only 0.78 s, while the running time of the traditional iterative algorithm was more than 20 s.

In 2020, the Shanghai Institute of Optics and Mechanics [133] developed blue-green laser diode radar equipment equipped with 486-nm and 523-nm wavebands, which were compatible with the detection needs of nearshore and coastal seawater. The experimental results showed that the 486-nm blue-green laser diode radar equipment could complete the 100-m communication experiment, while the 532-nm blue-green laser diode radar equipment could only complete the 75-m communication experiment.

In 2020, Wu [134] studied the design of an underwater-positioning communication system based on a photoelectric array. The system used light-emitting diodes as the emission light source, and the receiving end used a curved array detector composed of 32 photodiodes. It used quaternary PPM. The bit error rate was analyzed under different receiving angles. When the receiving angle was less than 50°, the bit error rate was 1.248×10^{-6}. When the receiving angle was greater than 50°, the bit error rate was 6.234×10^{-3}, resulting in accurate positioning and communication.

In 2020, Tong [135] used a 470-nm blue-green LED as the emission light source and an avalanche photodiode as the detector at the receiving end. Using OFDM, he completed an underwater communication experiment with a communication rate of up to 10 Mbit/s and a bit error rate as low as 10^{-4} in a 1.5-m pool in the laboratory.

In 2021, Dong et al. [136] used four 470-nm blue light-emitting diodes to form a light source array. The receiving optical transceiver used Fresnel lens imaging to

achieve large field-of-view reception. The detector used avalanche photodiodes. An underwater communication experiment was completed in a 20-m pool with a communication rate as high as 5 Gbit/s and a bit error rate as low as 10^{-6}.

In 2021, Xu et al. [137] proposed the use of soft spread spectrum software with five times the spread spectrum gain to improve the receiving sensitivity of underwater optical communication systems. Simulations verified that spread spectrum technology could extend the communication distance. The experimental results showed that when the signal bandwidth was 200 MHz, the sensitivity of the system receiver increased by 4.8 dB to −32.5 dBm, which is equivalent to the communication distance being extended by 5.8 m.

In 2021, He et al. [138] studied an underwater optical communication system with an adaptive alignment. The system used a high-power blue laser diode as the emission light source, and the receiver used an avalanche photodiode as the detector. An error-free underwater experiment was conducted in a pool with a 30-m communication distance and achieved a 100 Mbit/s transmission.

In 2021, Ma et al. [139] designed a high-speed underwater video transmission system using a novel underwater image transmission method. The system used a blue-green LED as the emission light source and PPM mode to complete the high-definition video transmission experiment. The communication distance was 1.4 m and the bit error rate was less than 10^{-6}.

In 2021, Li et al. [140] studied a DC-biased optical OFDM underwater optical communication system using digital preequalization technology. The system used a blue-green LED as the emission light source and a photodiode as the detector. An underwater experiment with a communication distance of 1 m, a data transmission rate of 32 Mbit/s, and a bit error rate lower than 10^{-3} was successfully completed in clean water.

In 2021, Li et al. [141] studied a MIMO underwater optical communication system. The system used six blue-green LED arrays as the emission light source, and the receiver detector was composed of three photomultiplier tubes. The data transmission rate was 1 Mbit/s, and the bit error rate was 1.1×10^{-5} in a 10-m-deep water tank in a laboratory. The signal-to-noise ratio was 21.4 dB.

In 2021, Chai [142] used a 470-nm blue laser diode as the emission light source, the receiver used an avalanche photodiode as the detector, and the improved OFDM mode was adopted. In a laboratory water tank, the data transmission rate was up to 6.28 Mbit/s, the communication distance was up to 1.8 m, and the bit error rate was as low as 9.552×10^{-6}.

In 2021, Shen [143] used a 24-W high-power blue-green LED as the transmitting light source, and the receiving optical antenna used an aspheric lens with a 50-mm diameter, and the switch keying modulation mode was adopted. An underwater experiment was completed with a data transmission rate of up to 18 Mbit/s, communication distance of up to 16 m, and communication error rate of less than 10^{-7}. It completely transmitted pictures, files, text, and other data streams.

In 2022, Fei et al. [144] used 450-nm laser diode and switch keying modulation. In a laboratory water tank, the data transmission rate was up to 3 Gbit/s, the

communication distance was up to 100.6 m, and the bit error rate was as low as 4.2×10^{-5}.

In 2022, Ke et al. at Xi'an University of Technology [145] designed an underwater optical communication model with a receiving antenna consisting of a double compound eye lens using a blue-green LED light source and a switch keying modulation mode. The system was tested in the laboratory through different suspended particle turbidities and transmission distances, verifying its feasibility.

In 2022, Shi et al. [146] studied a method to directly detect the channel of an underwater optical communication system by means of end-to-end deep learning and used this method to study and compare the signal-to-noise ratio and bit error rate of the underwater communication system. The experimental results showed that the depth learning algorithm performed better than the traditional least mean-squares algorithm. When the communication distance was 15 m, the SNR was 67 dB, and the bit error rate was 7.02×10^{-4}.

In 2022, Li et al. [147] studied a high-power underwater laser emission system device. The light source of the system used a 532-nm green laser diode with an average power of up to 2 W in an underwater optical communication experiment. The experimental results showed that when using a high-power laser diode, the communication distance was up to 150 m, the communication rate was up to 1.2 Mbit/s, and the bit error rate was 3.5×10^{-4}.

In 2022, Zhang et al. [148] studied an underwater weak light signal detection method based on adaptive stochastic resonance. By analyzing the characteristics of underwater weak light signals and the stochastic resonance detection mechanism, a second sampling was introduced to make stochastic resonance applicable to optical signal detection at any frequency. To verify the effectiveness of this method, an underwater optical wireless communication (UOWC) experimental system based on an avalanche photodiode was built. The experimental results showed that when the SNR was −4.559 dB, the bit error rate was 5×10^{-4}. This further proved that underwater weak light signal detection based on adaptive stochastic resonance could significantly improve the bit error rate performance of the UOWC receiver.

In 2022, Wang [149] used a blue-green laser diode as the emission light source, an avalanche photodiode as the detector at the receiving end, and the RS (255, 233) coding method. In the laboratory's water tank, he completed an underwater full-duplex communication experiment with a data transmission rate of up to 50 Mbit/s, a communication distance of up to 30 m, and a bit error rate as low as 10^{-13}.

In 2022, Chen et al. [150] used an LED as the emission light source and an avalanche photodiode as the detector at the receiving end. They successfully implemented an underwater communication experiment with a 2.5-m communication distance and a data transmission rate of up to 50 Mbit/s using the OFDM QAM method. The measured channel signal-to-noise ratio was 6.62 dB on average, and the maximum was 33.55 dB.

Domestic research progress on underwater optical communication systems is summarized in Table 31.4.

Table 31.5 summarizes the domestic research progress in underwater optical communication systems, according to the communication rates (kbit/s, Mbit/s, and Gbit/s).

Table 31.4 Domestic research progress in underwater optical communication systems

Year	Researcher	Research contents	Research results
1989	Ao et al.	Laser submarine communication using error simulation, pulse position modulation, and RS coding	RS error correction coding can greatly improve the channel bit error rate
1995	Wang et al.	Design of a new time slot synchronizer for laser underwater communication	The synchronizer has good anti-noise performance and can meet the requirements of an underwater signal detection and demodulation unit
1992	Sui et al.	Different modulation methods were used in an underwater communication system	When the relative phase shift keying modulation mode is used, the bit error rate is the minimum at 10^{-8}
1997	Yang et al.	Laser underwater communication system using photomultiplier tubes and RS coding	If the bit error rate is controlled below 10^{-8}, more than 1 dB of coding gain can be obtained
1999	Ao et al.	Underwater laser communication model using an adaptive array intensifier and diversity reception	The output signal-to-noise ratio of the adaptive array signal intensifier is proportional to the number of array input signals
2006	Ao et al.	A highly sensitive digital optical receiver for underwater communication systems was designed	The receiver reduces the system requirements for the signal-to-noise ratio and achieves good performance
2012	Wang et al.	Variable gain amplifier was used to reduce the dynamic range of the received signal, and two controllable gain methods were proposed	This method can effectively compress the dynamic range of the signal and make it stable
2021	Xu et al.	Used soft spread spectrum software with five times the spread spectrum gain to improve the receiver sensitivity of underwater optical communication systems	The sensitivity of the receiving end of the system increased to −32.5 dBm, 4.8 dB higher, which is equivalent to extending the communication distance 5.8 m
2022	Ke et al.	Designed an underwater optical communication model with a blue-green LED light source, and a double compound eye lens composed of a receiving antenna and a compound eye lens	The communication distances are 1 m, 3 m, and 5 m. the received optical power of the double-layer compound eye system is, respectively, 72%, 65%, and 60% higher than that of the single-layered system

Table 31.5 Domestic progress in terms of kbit/s, Mbit/s, and Gbit/s communication rates

Classification	Year	Researcher	Communication rate	Communication distance	Error rate
Kbit/s	1987	Jiang et al.	640 kbit/s	–	5.2×10^{-2}
	1998	Huang et al.	21.2 kbit/s	3 m	2.8×10^{-3}
	2002	Huang et al.	20 kbit/s	50 m	1.76×10^{-3}
	2008	Luo et al.	125 kbit/s	2 m	Error free
	2009	He et al.	30 kbit/s	–	Error free
	2010	Liu et al.	–	60 m	$< 10^{-4}$
	2013	Zhu et al.	73 kbit/s	2 m	Error free
	2013	Tian et al.	38.4 kbit/s	30 m	–
	2015	Cai et al.	115.2 kbit/s	10 m	Almost error free
	2018	Zhang	9.6 kbit/s	3 m	$< 10^{-5}$
Mbit/s	2004	Pang	1/3 Mbit/s	10/13 m	8×10^{-9}/ 7.31×10^{-9}
	2008	Cui	2 Mbit/s	30 m	–
	2011	Cong	3 Mbit/s	20 m	4×10^{-4}
	2013	Xu et al.	161.36/156.31/ 131.07 Mbit/s	–	2.5×10^{-3}/7.42×10^{-4}/3.17×10^{-3}
	2016	Hu et al.	60 Mbit/s	5 m	Error free
	2017	Ao et al.	10 Mbit/s	2.3 m	Error free
	2017	Lin	80 Mbit/s	2 m	2.24×10^{-3}
	2017	Song	31.082 Mbit/s	2 m	$< 10^{-4}$
	2018	Wang et al.	0.98 Mbit/s	6 m	Error free
	2018	Han	20 Mbit/s	9.6 m	10^{-9}
	2018	Yang et al.	6 Mbit/s	3 m	7×10^{-4}
	2018	Li	100 Mbit/s	16 m	2.4×10^{-6}
	2018	Wang et al.	25 Mbit/s	10 m	$< 10^{-4}$
	2019	Wang et al.	500 Mbit/s	100 m	2.5×10^{-3}
	2019	Wang et al.	50 Mbit/s	3 m	7.11×10^{-4}
	2019	Shang	75 Mbit/s	4 m	Error free
	2020	Li et al.	1 Mbit/s	10 m	3.9×10^{-6}
	2020	Tong	100 Mbit/s	1.5 m	$< 10^{-4}$
	2021	Li et al.	32 Mbit/s	1 m	10^{-3}
	2021	Li et al.	1 Mbit/s	10 m	1.1×10^{-5}
	2021	He et al.	100 Mbit/s	30 m	Error free
	2021	Chai	6.28 Mbit/s	1.8 m	$< 9.552 \times 10^{-6}$
	2021	Shen	18 Mbit/s	16 m	$< 10^{-7}$
	2022	Wang	50 Mbit/s	30 m	$< 10^{-13}$
	2022	Chen et al.	50 Mbit/s	2.5 m	–
	2022	Li et al.	1.2 Mbit/s	150 m	3.5×10^{-4}
Gbit/s	2015	Liu et al.	2.7 Gbit/s	35 m	3.8×10^{-3}
	2016	Shen et al.	2/1.5 Gbit/s	12/20 m	2.8×10^{-5}/3.0×10^{-3}

(continued)

Table 31.5 (continued)

Classification	Year	Researcher	Communication rate	Communication distance	Error rate
	2016	Lu et al.	9.6 Gbit/s	8 m	10^{-4}
	2016	Xu et al.	1.324/4.883 Gbit/s	6 m	2.02×10^{-3}/3.20×10^{-3}
	2017	Chen et al.	5.3/5.5 Gbit/s	21 m	2.47×10^{-3}/2.92×10^{-3}
	2018	Huang et al.	14.8/10.8 Gbit/s	1.7/10.2 m	1.9×10^{-5}/1.5×10^{-4}
	2019	Hong et al.	18.09/19.21/ 12.62 Gbit/s	5/25/35 m	Capacity increased by 32.22%, 30.03%, and 31.55%, respectively
	2019	Zhou et al.	15.17 Gbit/s	1.2 m	3.8×10^{-3}
	2019	Han	6.23 Gbit/s	66 m	$< 10^{-4}$
	2020	Ji et al.	5/4.5 Gbit/s	50/55 m	–
	2021	Dong et al.	5 Gbit/s	20 m	10^{-6}
	2022	Fei et al.	3 Gbit/s	100.6 m	4.2×10^{-5}

With the improvement in channel coding and modulation, the communication rates of systems are increasing from kbit/s and Mbit/s to Gbit/s. However, in a real underwater environment, the channel is subject to the influences of temperature, salinity, and other parameters, and most of the above experiments were carried out in laboratories, not in real rivers and oceans. This may result in a significant difference between the experimental results and the actual application results. Therefore, actual ocean exploration needs further research.

Moreover, the communication distance of most underwater optical wireless communication system experiments can only reach a few meters to tens of meters, and only a few experiments reach 100 m. Therefore, increasing the communication distances of underwater optical wireless communication systems is a current challenge, and more experiments are required to address this in the future.

Key Technologies in Underwater Optical Wireless Communication

Channel Modeling of Underwater Optical Wireless Communication Systems

Optical Properties of Water

According to Mobley's theory, water has two types of optical properties: intrinsic and apparent [11]. The inherent optical properties of water mainly include absorption and scattering coefficients. Compared with the apparent optical properties, the

inherent optical properties have a greater impact on the performance of underwater optical wireless communication; therefore, research has mainly focused on inherent optical properties.

The underwater environment is complex, and absorption and scattering have a serious impact on beam propagation, resulting in attenuation [151]. The absorption characteristic shortens the communication distance, and the scattering characteristic affects the propagation direction of light. This results in a multipath effect that reduces the signal-to-noise ratio and increases the bit error rate. Figure 31.2 shows the geometric structure of the inherent optical properties of seawater.

According to the law of conservation of energy [37],

$$P_I(\lambda) = P_A(\lambda) + P_S(\lambda) + P_T(\lambda), \tag{31.1}$$

where λ is the wavelength, ΔD is the width of the water tank, $P_I(\lambda)$ is the incident light power, $P_A(\lambda)$ is the absorbed light power of the water, $P_S(\lambda)$ is the scattered light power, and $P_T(\lambda)$ is the output light power.

Using the absorbance definition formula, the absorption coefficient can be obtained as follows [37]:

$$a(\lambda) = \lim_{\Delta D \to 0} \frac{P_A}{P_I \Delta D}. \tag{31.2}$$

The underwater light absorption coefficient can be expressed as four parts: pure seawater absorption ($a_\omega(\lambda)$), detritus absorption ($a_{\text{det}}(\lambda)$), phytoplankton absorption ($a_{phy}(\lambda)$), and colored dissolved organic matter absorption ($a_{CDOM}(\lambda)$). Therefore [1],

$$a(\lambda) = a_\omega(\lambda) + a_{\text{det}}(\lambda) + a_{phy}(\lambda) + a_{CDOM}(\lambda). \tag{31.3}$$

Using the definition for the scattering degree, the scattering coefficient can be obtained as follows [37]:

$$b(\lambda) = \lim_{\Delta D \to 0} \frac{P_S}{P_I \Delta D}. \tag{31.4}$$

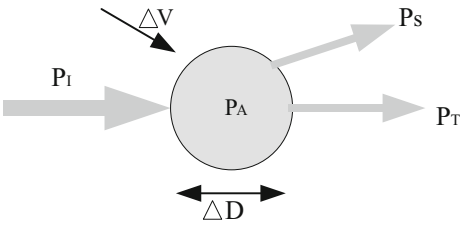


Fig. 31.2 Geometric structure of inherent optical properties of seawater

The underwater light scattering coefficient can be represented by three parts: pure seawater scattering ($b_\omega(\lambda)$), detritus scattering ($b_{\text{det}}(\lambda)$), and phytoplankton scattering ($b_{phy}(\lambda)$); hence [1],

$$b(\lambda) = b_\omega(\lambda) + b_{\text{det}}(\lambda) + b_{phy}(\lambda). \tag{31.5}$$

The attenuation coefficient of light $c(\lambda)$ can be expressed as [37].

$$c(\lambda) = a(\lambda) + b(\lambda), \tag{31.6}$$

where the attenuation coefficient is expressed in m^{-1}. The attenuation coefficient $c(\lambda)$ is related to many factors of the seawater environment, including the seawater salinity, turbidity, and even plankton.

The correlation coefficients for different seawater samples are listed in Table 31.6. It can be seen that the main substances in pure seawater are water and inorganic salts; thus, the attenuation is mainly caused by absorption. The scattering coefficient is relatively low and can be ignored. In clear seawater, the concentration of colored dissolved organic matter is relatively large, which mainly affects the light scattering. In coastal seawater, phytoplankton and debris cause light absorption and scattering. In turbid harbors, suspended particles and debris cause severe light beam attenuation.

The Beer–Lambert law can be used to describe the attenuation effect of light under water [152]; its expression is as follows:

$$I = I_0 e^{-c(\lambda)d}, \tag{31.7}$$

where I_0 is the transmitted light power, $c(\lambda)$ is the attenuation coefficient, d is the transmission distance, and I is the residual optical power after reaching the transmission distance, d. The value of $c(\lambda)$ changes with the difference in the water area type and water depth.

A model with two exponential functions that can more accurately describe the power loss of the underwater channel can also be obtained from the Beer–Lambert law; its expression is as follows [153]:

$$I = I_0 e^{-c_1(\lambda)d} + I_1 e^{-c_2(\lambda)d}. \tag{31.8}$$

In this formula, the first index indicates that the diffusion length is greater than the attenuation loss length, and the second index indicates that the diffusion length is less than the attenuation loss length. The specific parameters are listed in Table 31.7.

Table 31.6 Typical absorption ($a(\lambda)$), scattering ($b(\lambda)$), and attenuation ($c(\lambda)$) values for four different water qualities [58]

Water quality	a(λ)(m^{-1})	b(λ)(m^{-1})	c(λ)(m^{-1})
Pure seawater	0.041	0.003	0.044
Clear seawater	0.114	0.037	0.151
Coastal seawater	0.179	0.219	0.398
Turbid port water	0.366	1.824	2.190

Table 31.7 Parameters for the power loss model of a long-distance underwater channel [154]

Seawater conditions	$C(mg/m^3)$	$I_0(m^{-1})$	$c_1(\lambda)(m^{-1})$	$I_1(m^{-1})$	$c_2(\lambda)(m^{-1})$
Pure seawater	0.0050	0.2000	0.0657	0.0046	0.2634
Clear seawater	0.3100	0.1000	0.1508	0.1589	0.4937

Generally speaking, when light travels underwater, absorption and scattering cause three problems: First, absorption reduces the total energy of the light, which limits the communication distance. Second, scattering diffuses the beam, thereby reducing the total number of photons received. Third, scattering causes photons to arrive at the receiver at different times, resulting in a multipath phenomenon. Therefore, to solve these problems, extensive research and experiments should be conducted on the factors that affect underwater channels. Over the past few years, researchers have designed many reliable underwater optical wireless communication system models. The following are the three main underwater-channel modeling methods.

Model of an Underwater Turbulent Channel

When light propagates underwater, it is affected not only by absorption and scattering but also by underwater turbulence. Underwater turbulence refers to a change in the random refractive index when a light beam passes through a water body [155]. The refractive index of underwater turbulence is not only affected by temperature but also by the water quality, salinity, and other factors. Underwater turbulence also changes with a change in geographical location [156]. However, it still conforms to the hydrodynamic equation [157].

Nikishov [158] proposed a method to describe the complex underwater turbulence spectrum, considering the temperature, salinity, refractive index, and other factors, which well simulate the transmission of light in water. The underwater refractive index fluctuation spectrum can be expressed as [159]

$$\begin{aligned}\Phi(k) = {} & 0.388 \times 10^{-8}\varepsilon^{-1/3}(k)^{-11/3}\left[1 + 2.35(\eta k)^{2/3}\right] \\ & \times \frac{\chi_T}{w^2}\left(w^2 e^{-A_T\delta} + e^{-A_S\delta} - 2we^{-A_{TS}\delta}\right),\end{aligned} \tag{31.9}$$

where k is the spatial wave number, ε is the kinetic energy dissipation rate of the water medium, η is the Kolmogorov internal scale, and χ_T is the temperature dissipation rate. From the water surface to underwater, the range of χ_T is [10] K^2/s. w is the underwater turbulence ratio caused by salinity and temperature, $A_T = 1.863 \times 10^{-2}$, $\delta = 8.284(\eta k)^{4/3} + 12.987(\eta k)^2$, $A_S = 1.9 \times 10^{-4}$, and $A_{TS} = 9.41 \times 10^{-3}$.

Because the underwater environment is more complex than the atmospheric environment, multiple underwater parameters must be expressed together. The following are the four turbulence parameters of the underwater turbulence intensity.

1. Temperature Dissipation Rate χ_T: The temperature change is closely related to the turbulence state. The temperature dissipation rate can be used to express the influence of the temperature change on the underwater turbulence parameters; its expression is [160]

$$\chi_T = 2\kappa_T \left\langle \left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2 \right\rangle, \tag{31.10}$$

where T represents the temperature value and κ_T is the temperature difference coefficient. Assuming that the water quality is uniform, $\frac{\partial T}{\partial x}$ represents the temperature gradient in the x direction, $\frac{\partial T}{\partial y}$ represents the temperature gradient in the y direction, and $\frac{\partial T}{\partial z}$ represents the temperature gradient in the z-direction. The temperature gradient at each location is the same. The above equation can be simplified as [161]

$$\chi_T = 6\kappa_T \left\langle \left(\frac{\partial T}{\partial z}\right)^2 \right\rangle, \tag{31.11}$$

where T represents the temperature, κ_T is the temperature difference coefficient, and $\frac{\partial T}{\partial z}$ represents the temperature gradient in the z-direction. The larger the temperature dissipation rate, the stronger the stratification of the water area, and the stronger the underwater turbulence.

2. Kinetic Energy Dissipation Rate ε: With isotropy and uniform water quality, the kinetic energy dissipation rate per unit of fluid mass can be expressed as [161]

$$\varepsilon = \frac{15v\left\langle (\partial u/\partial z)^2 \right\rangle}{2}, \tag{31.12}$$

where v is the viscosity of the molecular motion, and $\partial u/\partial z$ represents the derivative of the turbulent velocity.

3. Underwater Turbulence Ratio w Caused by Temperature and Salinity: When the temperature gradient is constant, w can be expressed as [161]

$$w = \frac{\alpha(\Delta T/\partial z)}{\beta(\Delta S/\partial z)}, \tag{31.13}$$

where α and β are constants, ΔT is the temperature difference between the water surface and underwater, and ΔS is the salinity difference between the water surface and underwater.

4. Kolmogorov Internal Scale η: The minimum underwater turbulence scale is the Kolmogorov internal scale, and its expression is [158]

$$\eta = \left(v^3/\varepsilon\right)^{1/4}, \tag{31.14}$$

where v is the kinematic viscosity variable, and $\varepsilon$is the turbulent flow energy dissipation rate. From Eq. 31.14, it can be seen that the Kolmogorov internal scale increases with an increase in the kinematic viscosity and decreases with an increase in the turbulent flow energy dissipation rate.

In recent years, many researchers have investigated the effects of turbulence on underwater optical wireless communication systems and also proposed models to suppress the turbulence effects. Marine turbulence channel models include the generalized gamma turbulence model, logarithmic normal turbulence model [159], and the MIMO channel turbulence model [162].

In 2014, Hamza et al. [163] established a fading model of a Gaussian lognormal distribution in weak ocean turbulence using the Rytov method. A formula for the on-axis scintillation index of focused Gaussian beams in weak ocean turbulence was derived, and the average bit error rate was calculated using this formula. In the same year, Tang et al. [164] used a double-gamma function to model the impulse response of an underwater optical communication link in coastal and port waters. The model was consistent with the Monte Carlo simulation results, and the effective region of the model was further studied.

In 2017, Oubei et al. [73] modeled an underwater channel with a temperature gradient and adopted an experimental method with a generalized gamma distribution. The experimental results showed that the measured data under all underwater channel conditions fit the distribution. In addition, it was observed that a thermally homogeneous channel could be well described by a simple gamma distribution. This was the first model to comprehensively describe a thermally homogeneous and gradient-based underwater optical channel.

In 2020, Du [165] studied the bit error rate performance of optical wireless communication systems using pulse position modulation (PPM) in ocean anisotropic turbulence and calculated the bit error rate using the double-gamma distribution method. In 2021, He et al. [166] experimented with underwater optical communication systems using a marine turbulent channel transmission model with a double-gamma distribution. They analyzed and compared the average bit error rate of underwater optical communication systems with different wavelengths.

Jamali et al. [74] used MIMO spatial diversity technology on an underwater optical wireless communication link to mitigate the fading caused by turbulence. Considering the effects of absorption and scattering, the exact expressions of the optimal gain, equal gain combination, and upper bound of the BER were obtained using the analytical solution method.

In 2016, Li et al. [167] considered anisotropic turbulence using a Gaussian–Schell model of a seawater channel. It was concluded that the greater the temperature fluctuation was, the greater the depolarization effect. However, the depolarization effect of anisotropic turbulence was weaker than that of isotropic turbulence. In 2019, Zou et al. [168] used a partially coherent Gaussian beam SISO model to

analyze the average capacity of an underwater optical wireless communication system under weak ocean turbulence.

In 2021, Jiang [159] analyzed different ocean turbulence models, including the lognormal distribution, exponential logarithmic distribution, Weibull distribution, gamma–gamma distribution, and generalized gamma distribution, and provided their probability density functions. Finally, the Monte Carlo simulation method was used to verify the accuracy of asymmetric clipped optical OFDM and DC-biased optical OFDM spatial diversity systems under different turbulence environments.

At present, the characteristics of underwater channels in underwater optical communication systems are mainly studied in the laboratory through simulations. The simulated marine environment is somewhat different from actual seawater, which makes the experimental data insufficient and reliable. In the future, a comprehensive experimental study of the influence of turbulence on underwater channels should be conducted in a real marine environment.

Channel Attenuation Model

Because the underwater environment is much more complex than the air environment, underwater optical wireless communication research is challenging, owing to the absorption and scattering of light when it propagates underwater. The attenuation of light beams in water is approximately 100–1000 times higher than that in air [169]. Therefore, one of the main objectives of underwater optical communication channel modeling is to evaluate the overall path attenuation loss.

The optical path attenuation model of underwater optical communication mainly consists of direct and indirect view models, both of which can be obtained using the Monte Carlo method. These modeling methods are described in detail below.

Direct Viewing Model

Two methods are used for modeling the direct view model: the radiative transfer equation and the Beer–Lambert law. The light attenuation effect in a simple underwater environment can be expressed using the Beer–Lambert law [1]. The Beer–Lambert law is an exponential attenuation model that is typically used to calculate the optical path loss [170]. Smar [170] used the Beer–Lambert law to study and experiment with a direct viewing underwater optical communication system and obtained the performance of an underwater optical wireless communication system under different water qualities and communication distances.

However, in actual underwater communication, most researchers have chosen the radiative transfer equation model. The radiative transfer equation is a general theoretical model. Jaruwatanadilok expressed the radiative transfer equation as [56]

$$\vec{n}\cdot\nabla L\left(\lambda,\vec{r},\vec{n}\right)=-cL\left(\lambda,\vec{r},\vec{n}\right)+\int_{2\pi}\beta\left(\lambda,\vec{n},\vec{n}'\right)L\left(\lambda,\vec{r},\vec{n}'\right)d\vec{n}'+E\left(\lambda,\vec{r},\vec{n}\right). \tag{31.15}$$

All parameters in the formula are related to the wavelength. c is the attenuation coefficient, ∇is the divergence operator, $L\left(\lambda, \vec{r}, \vec{n}\right)$ is the optical radiance propagating in direction $\vec{n}$ at location $\vec{r}$, $E\left(\lambda, \vec{r}, \vec{n}\right)$ is the source radiance, and $\beta\left(\lambda, \vec{n}, \vec{n}'\right)$ is the volume scattering function.

The radiative transfer equation model includes two common solutions: analytical and numerical [171]. The radiative transfer equation considers the polarization and multiple scattering of light, which can more accurately simulate the attenuation of underwater beam propagation. However, the radiative transfer equation is a complex multivariable calculus equation, and it is very difficult to find a general analytical solution form. Therefore, researchers generally use the probability method to obtain a numerical solution to the radiative transfer equation.

Numerical solution methods include the Monte Carlo, discrete ordinate, random model, and invariant embedding methods [14]. Because the invariant embedding method can only solve a one-dimensional radiative transfer equation, it is unsuitable for underwater optical wireless communication systems. The discrete ordinate method is difficult to program and generally not selected [171]. The stochastic model, based on the probability property of the photon trajectory, is not yet mature and therefore rarely used. The Monte Carlo method is the most widely used. Therefore, only the application of the Monte Carlo method to underwater optical wireless communication systems is studied here.

In 2008, Jaruwatanadilok [56] used the vector radiative transfer theory to model an underwater optical communication system and quantized the relationship between the scattering, distance, and bit rate through a Monte Carlo simulation. In 2009, Cochenour et al. [58] proposed a beam spread function for an underwater optical communication system and used the small-angle approximation method to solve the radiative transfer equation.

In 2015, Li et al. [171] proposed a fast numerical solution for the radiative transfer equation to calculate the optical path loss of underwater channel attenuation. The results showed that this method was faster than the traditional Monte Carlo method. In 2018, Illi et al. [76] proposed an improved radiative transfer equation to predict the optical path loss in an underwater channel and calculated the change in the bit error rate performance of an underwater optical wireless communication system with the transmission time and distance.

In 2019, Illi et al. [77] further improved the radiative transfer equation calculation method. The experimental results showed that the improved radiative transfer equation could obtain the same results by calculating the bit error rate of the underwater optical wireless communication system. It was consistent with the Monte Carlo simulation results, thereby reducing the complexity of the system.

Because of the complexity of calculating the radiative transfer equation, it is difficult to obtain an accurate solution. The Monte Carlo method is simple to program, flexible, and calculates accurately; therefore, it is often used to solve the radiative transfer equation. Using Monte Carlo simulations, the US Naval Laboratory published a usage guide for underwater optical wireless communication. It was

proven effective by many researchers and adopted by many scholars studying underwater optical wireless communication [1].

Li et al. [172] used the Monte Carlo model to study the channel capacity of underwater optical communications. This model can flexibly represent various parameters of the transmitter, receiver, and channel and is applicable to the design of physical underwater optical communication links. Gabriel et al. [173] proposed the channel impulse response of underwater optical wireless communication systems for different water quality types by solving the radiative transfer equation through a Monte Carlo simulation. The above research methods are mainly aimed at the underwater optical communication channel model, and there is no requirement for the communication distance.

Hanson [54] conducted an experiment by simulating dozens of meters of seawater path using the Monte Carlo method; the results showed that the underwater optical communication rate reached 1 Gbit/s. Dalgleish et al. [174] used the Monte Carlo method to calculate a 500-ps laser pulse in a 12.5-m test box and obtained the pulse response under different environments.

In addition to using the Monte Carlo method, Ding et al. [175] introduced a multiple scattering channel model, in which the optical power at the receiver could be solved using Monte Carlo integration as the multiple integration of the probability density function.

In 2020, Yuan et al. [176] obtained the impact response function of the system through Monte Carlo integration in optical wireless multiscattering channel communication. The experimental results showed that for third-order scattering, compared with the original Monte Carlo integration based on uniform sampling, the computational efficiency of the Monte Carlo integration model based on partial importance sampling was improved by 12 times, which is 5.6 times that of the widely used Monte Carlo simulation model. For high-order scattering, the Monte Carlo integration model based on partial importance sampling has higher computational efficiency than the Monte Carlo simulation model.

Owing to the simplicity of modeling and implementing direct viewing channels, most underwater optical wireless communication experimental systems use direct viewing links. However, the direct view model must be a point-to-point transmission, and it is difficult to achieve point-to-point transmissions in underwater environments, owing to the scattering effect; therefore, the direct view model has certain limitations in underwater optical wireless communication systems.

Indirect Viewing Model

In traditional direct view links, signals are transmitted from point to point. Obstacles in the propagation path can attenuate or even hinder the propagation of optical signals. An indirect viewing model was proposed to solve this problem. The indirect viewing model uses a reflected beam of water molecules to overcome obstacles in the communication link [14]. Similar to the direct viewing channel model, the indirect viewing channel model is mostly simulated using the Monte Carlo method.

Arnon et al. [177] studied a new indirect viewing underwater optical wireless communication network using the Monte Carlo method and deduced the

mathematical model of the channel. Tang et al. [178] studied the path loss of an indirect viewing underwater optical wireless communication link, based on Monte Carlo simulation, and considered the influence of seawater scattering characteristics. The experimental results showed that with an increase in attenuation length, the scattered light could eliminate the influence of the random surface slope caused by wind on the received signal.

Jagadeesh et al. [179] proposed an indirect viewing underwater optical wireless communication model to evaluate channel characteristics using the Monte Carlo simulation method. The performance of the underwater system under different water types and receiving fields of view was characterized and simulated using the derived impulse response model. The results showed that this model provided a new method for characterizing indirect viewing underwater optical wireless communication channels.

Anous et al. [180] evaluated the system performance of direct and indirect view underwater optical wireless communication, based on the underwater optical communication vertical link model. Considering the heterogeneity of underwater environments, a simple two-layer model was established to accurately calculate the link budget. Doniec et al. [64] introduced a general model of an end-to-end underwater optical communication system, which provides a reference for later research on underwater communication models.

Modeling Cross-Media Transmissions

In an underwater optical communication system, the light beam must cross atmospheric, air, sea, and water channels. Different channels differently affect the optical transmission characteristics. The substances that affect the beam propagation in atmospheric channels include aerosols, atmospheric molecules, fog, and rain. Because the sea surface is affected by the sea wind, the light beam experiences serious refraction and reflection. The main factors affecting light beam transmission in water channels are biological particles and soluble organic matter.

At present, extensive research has been conducted on the transmission characteristics of light beams in different channels at home and abroad; however, most of them are for a single channel, which cannot meet the design requirements of cross-medium channel modeling.

The wave spectrum selected for wave modeling is the Pierson–Moskowitz (PM) spectrum. Its specific expression is [181]

$$S(\omega) = \frac{g^2\kappa}{\omega^5} \exp\left[-\vartheta\left(\frac{g}{U\omega}\right)^4\right], \tag{31.16}$$

where g is the acceleration of gravity, ω is the angular frequency, $\kappa = 8.1 \times 10^{-3}$ is the dimensionless quantity, $\vartheta = 0.74$, and U is the wind speed 21.5 m from the sea surface.

The wave spectrum is generally the product of the spectrum and direction function, and its expression is [181]

$$S(\omega, \theta) = S(\omega)\varphi(\theta), \tag{31.17}$$

where $S(\omega)$ is the frequency spectrum, $\varphi(\theta)$ is the direction function, and θ is the wind direction angle in the propagation direction.

The direction function is more specific than the frequency spectrum when describing the propagation direction of the sine wave, and its expression is [182]

$$\varphi(\theta) = \frac{2}{\pi}\cos^2(\theta). \tag{31.18}$$

When the laser is at the air–sea interface, the reflectivity and transmissivity can be calculated according to the polarization state of the laser. According to Fresnel's law and Snell's law, the total transmissivity of the air–sea interface is [183]

$$L_{bd} = L_{bd1} + L_{bd2} \tag{31.19}$$

$$L_{bd1} = \frac{1}{2}\left|\frac{\sin 2\alpha \sin 2\beta}{\sin(\alpha+\beta)} \cdot \frac{1 + \cos^2(\alpha-\beta)}{\cos^2(\alpha-\beta)}\right|, \tag{31.20}$$

where L_{bd1} is the transmissivity of the air–sea interface itself, α is the reflection angle, β is the transmission angle, and L_{bd2} is the transmissivity of bubbles and white waves passing through the air–sea interface.

Currently, cross-medium wireless communication between underwater and air has become an important development field. Liu et al. [182] studied the distribution and loss characteristics of a cross-medium uplink laser in the transmission process, using a Monte Carlo simulation method at different underwater transmission distances and different sea surface wind speeds. Li et al. [184] used the Monte Carlo simulation method to analyze the downlink cross-medium transmission characteristics of light in an atmospheric air–seawater channel under sunny, stratiform, and cirrus cloud weather conditions.

Si et al. [185] studied seawater–atmosphere cross-medium uplink communication and described in detail the impact of the wave surface on optical uplink transmissions. Chen et al. [186] discussed light scattering across the air–sea interface using the Kirchhoff model and first-order small perturbation model theory in rough surface scattering. Wang et al. [187] studied the factors that affect laser transmission in the air–seawater cross-medium channel and deduced the energy equation of blue-green light passing through the cross-medium channel.

Li et al. [188] studied the transmission characteristics of the air–underwater cross-medium channel and concluded that its attenuation was mainly related to weather conditions, seawater transparency, etc. Zhang et al. [189] established a model for the dynamic downward transmission of a blue-green laser across the air–sea channel by considering real seawater conditions, such as different near-sea atmospheric sea-fog visibility, chlorophyll concentrations, and bubble concentrations.

Research on cross-medium channels is still limited. For example, in experiments on vertical links, researchers regard seawater as a uniform medium; however, the salinity and chlorophyll concentration differ at different depths in vertical seawater,

and the experimental results will not match the actual situation. At present, most research on cross-media channels is about downlink transmissions, while research on uplink and downlink transmissions is scarce.

Light Source Selection

Different light sources have different transmission distances underwater, and the stability of the light path is also different. The choice of the light source can determine the communication distance and modulation bandwidth. Seawater absorbs less blue-green light in the 450–550 nm wavelength range [39]. Seawater near the coast is more suitable for using green light as the light source, and the deep sea is more suitable for using blue light as the light source [190].

There are three types of light sources for underwater optical communication: light-emitting diodes (LEDs), laser diodes, and superradiant laser diodes. LEDs have the advantages of low cost, convenient production, and good safety; therefore, LEDs are currently the most widely used light sources. However, LEDs have disadvantages of non-aggregation, poor coherence, a large divergence angle, and difficulty with underwater transmissions.

The laser diode has good directivity and coherence, but is sensitive to temperature; therefore, a temperature controller is needed for applications. Superradiant laser diodes are semiconductor optoelectronic devices between light-emitting diodes and laser diodes that can output weak coherent light [191]. However, their reliability has not yet been evaluated and there is a risk of poor coupling [192]. Therefore, Table 31.8 only compares the main performance of LEDs and laser diodes.

The light sources used in underwater optical wireless communication systems are mainly blue-green diodes or laser diodes with a waveband between 450 and 550 nm. The communication distance and rate differ for different light source selections, and the communication distance can be increased by selecting an appropriate light source.

Tian et al. [193] used a blue GaN micro-light-emitting diode as the light source and successfully achieved a data rate of 200 Mbit/s, a communication distance of 5.4 m, and a bit error rate of 3×10^{-6} in underwater experiments. Liu et al. [194]

Table 31.8 Main performance comparison of LEDs and laser diodes

Parameter	Light-emitting diode	Laser diode
Coherence	Incoherent	Coherent
Modulation bandwidth	MHz magnitude	GHz magnitude
Divergence angle	Wide	Narrow
Spectral width	20–100 nm	0.01–5 nm
Photoelectric efficiency	About 50%	About 30%
Service life	Long	Moderate
Temperature sensitivity	Insensitive	More sensitive
Security	Good	Poor

experimentally demonstrated the use of white light generated by hybrid red, green, and blue laser diodes in underwater communication systems. The experimental results showed that the data rate was up to 8.7 Gbit/s in the 2.3-m underwater link.

Chao et al. [195] used a 450-nm laser diode as a light source in an underwater communication system. The experimental results showed that when the communication distances were 5 m, 35 m, and 55 m, the data transmission rates reached 19.6 Gbit/s, 13.2 Gbit/s, and 6.6 Gbit/s, respectively. Chen et al. [196] completed an underwater experiment with a net data rate of 3.31 Gbit/s and communication distance of 56 m using a 520-nm laser diode and 32-OAM mode.

Arvanitakis et al. [80] used a micro-light-emitting diode array composed of six 80-μm light-emitting diodes in series; it operated at 450 nm. When the communication distance was 4.5 m, the data transmission rate was as high as 3.4 Gbit/s.

Modulation Technology

Underwater optical wireless communication is limited by the devices in the blue and green light bands; thus, it is still in the low communication rate stage. It mainly includes on–off keying (OOK) modulation, pulse position modulation (PPM), and quadrature amplitude modulation (QAM). These modulation methods can achieve high data transmission rates in underwater communication systems; however, the communication distance is quite short.

In recent years, many scholars have achieved good results using improved modulation methods, such as polarized PPM [197], QAM orthogonal frequency-division multiplexing (OFDM) [198], polarized differential pulse position modulation [199], binary polarization shift keying [200], and binary phase shift keying [36].

Through the simulation and comparison of different modulation modes in underwater optical communication systems, Su et al. [201] concluded that PPM was more suitable for low-power submarine systems, and phase shift keying had the best performance in terms of bandwidth and error performance; however, its power efficiency was low. Xi et al. [103] used 4-PPM in underwater optical communications. The experimental results showed that the bit error rate of PPM was almost the same as that of switch keying modulation, and it had a higher spectral efficiency, which proved that 8-PPM and 16-PPM had higher bandwidth utilization.

Gabriel et al. [202] compared several commonly used modulation technologies in underwater optical wireless communication systems, in terms of their energy and bandwidth efficiency. The results showed that although PPM was optimal in terms of energy efficiency, differential PPM had advantages in terms of communication performance and equipment complexity.

Coding Technology

The influences of absorption and scattering on underwater light transmissions not only reduce the communication distance but also increase the bit error rate. Choosing

an appropriate coding technology can reduce the bit error rate. The coding technology used in underwater communication is mainly convolutional codes and block codes [203]. Convolutional codes include low-density parity check codes and turbo codes [204], which include cyclic redundancy check codes, Reed–Solomon (RS) codes, and linear cyclic error correction codes [203].

The RS code is the most widely used coding method in underwater optical wireless communication systems. Cox et al. [205] used a 405-nm laser diode to transmit data to a photodiode receiver using the RS (255, 129) code at a 500-kbit/s communication rate. The experimental results showed that, compared with the uncoded switch keying modulation system, the power required by this coding method was reduced by approximately 8 dB.

Simpson et al. [59] achieved a communication rate of 5 Mbit/s in 3-m and 7.7-m long water tanks by using RS coding. The experiment proved that when the bit error rate was 10^{-6}, the RS (255, 129) code improved the SNR by approximately 6 dB and the RS (255, 223) code improved the SNR by approximately 4 dB.

Cossu et al. [63] conducted an underwater experiment with a communication rate of 58 Mbit/s in a 3.3-m water tank using the RS-forward error correction (FEC) (272, 255) code. The bit error rate decreased from 1.5×10^{-3} to 10^{-9}. A few methods combine RS code and other coding methods. Doniec et al. [206] used a two-layer digital coding scheme that combined cyclic redundancy check codes and RS codes to transmit videos in an underwater optical wireless communication system. This multilayer coding scheme achieved low-delay transmission of video images underwater; however, the system was relatively complex.

Detection Technology

The detectors in underwater systems mainly include PIN photodiodes, avalanche photodiodes, and photomultiplier tubes. PIN photodiodes are mainly affected by thermal noise, whereas avalanche photodiodes are mainly affected by shot noise [207]. The PIN photodiode has a lower gain, whereas the avalanche photodiode has a higher gain. The communication distance in the underwater system is longer, but the internal circuit is more complex. Photomultiplier tubes have higher sensitivity, higher optical gain, and a lower noise level, but they are also have a high voltage supply (approximately 100 V) and high unit cost [208]. Their performances are compared in Table 31.9.

Chao et al. [195] used a silicon-doped PIN photodiode as a receiver to operate in the 320 nm–1100 nm wavelength range, with a response rate of approximately 0.15 A/W at 450 nm. Fei et al. [209] used silicon-doped avalanche photodiodes as receivers to achieve a data rate of 7.33 Gbit/s within 15 m. The avalanche photodiode operated in the wavelength range of 400 nm–1000 nm, and its response rate was approximately 4 A/W at 450 nm.

Sun et al. [210] used a 488-nm blue laser and a photomultiplier detector to achieve high-speed video transmissions with a transmission rate of 5 Mbit/s in a 4.5-m-long underwater channel. Dalgleish et al. [174] completed an underwater laser

Table 31.9 Detector performance comparison

Parameter	PIN photodiode	Avalanche photodiode	Photomultiplier tube
Volume	Small	Small	Large
Sensitivity	Low	Moderate	High
Detection area	Small	Small	Large
Equivalent noise power	High	Low	Extremely low
Use cost	Low	Moderate	High
Response speed	Slow	Fast	Ultrafast
Security	Good	Good	Low

measurement experiment of 500-ps laser pulse propagation using a photomultiplier detector at the receiving end in a 12.5-m water tank.

In underwater optical wireless communication systems, in addition to the three detectors described above, researchers have also doped with other chemicals to achieve better results. Hiskett et al. [66] used silicon avalanche photodiodes as detectors for underwater communication systems. An underwater experiment with a 40 Mbit/s transmission rate was completed in a 1-m-long water tank using a laser diode with a 450 nm wavelength. Li et al. [211] developed an underwater communication system, in which the receiver used a single-photon avalanche diode operating in Geiger mode as the detector, and realized long-distance communication.

Detectors that can detect weak underwater signals should be studied in the future. Detectors with high sensitivity, a large field of view, and low noise are future development directions. They are also a new challenge in underwater optical wireless communication systems.

Underwater Optical Wireless Communication System Design Model

To improve the optical power of the receiving end in an underwater optical wireless communication system, Ke et al. at Xi'an University of Technology studied an underwater communication system using wide-angle focusing double-layer fly eye lens optics [145]. The system used a 530-nm green LED as the emission light source of the underwater blue-green optical communication system and used a combination of a reflective cup and a focusing lens to focus and collimate the light source. In addition, a scheme for an LED drive circuit suitable for the transmitter of the underwater communication system was designed.

The receiving end adopted a compound eye lens optical receiving antenna with wide-angle focusing. According to the geometric optical principle of the lens array, the compound eye lens receiving system was designed and simulated using Zemax software. Finally, an experimental device for determining the receiving efficiency of an underwater communication system was developed.

As shown in Fig. 31.3, the communication performance of the system was tested with different turbidities and communication distances in the water tank and water

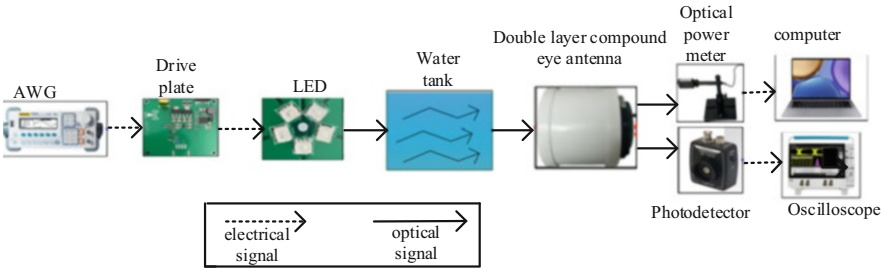


Fig. 31.3 Schematic diagram of the experimental device for reception efficiency

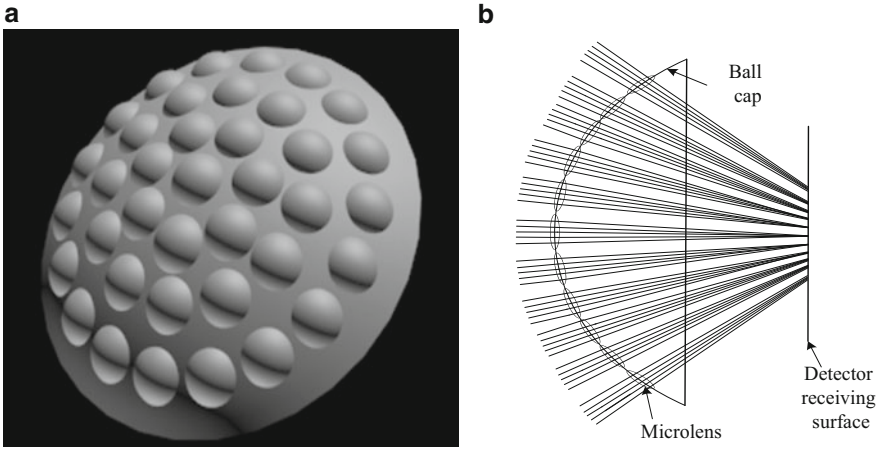


Fig. 31.4 Schematic diagram of an optical system with a single-layer curved compound eye lens array. (**a**) Single-layer curved compound eye lens structure. (**b**) Ray tracing of single-layer curved compound eye lens

pipe of the laboratory, respectively. The feasibility of this design scheme was verified by measuring the optical power of the receiving optical system.

The receiving end of the system adopted a compound eye lens optical receiving antenna with wide-angle focusing. Figure 31.4a shows the schematic of a single-layer curved compound eye lens. Because most of the detectors used in the laboratory are planar structures, the convergence effect of the microlens at the edge was reduced for the compound eye lens receiving antenna with a single-layer curved structure.

As shown in Fig. 31.4b, the lens at the center of the compound eye can focus the light beam on the surface of the detector; however, with the radial movement of the microlens on the surface of the spherical crown, the effect of the light beam focusing on the surface of the detector increasingly worsens, and the lens at the edge of the spherical crown can no longer focus the light beam on the surface of the photodetector.

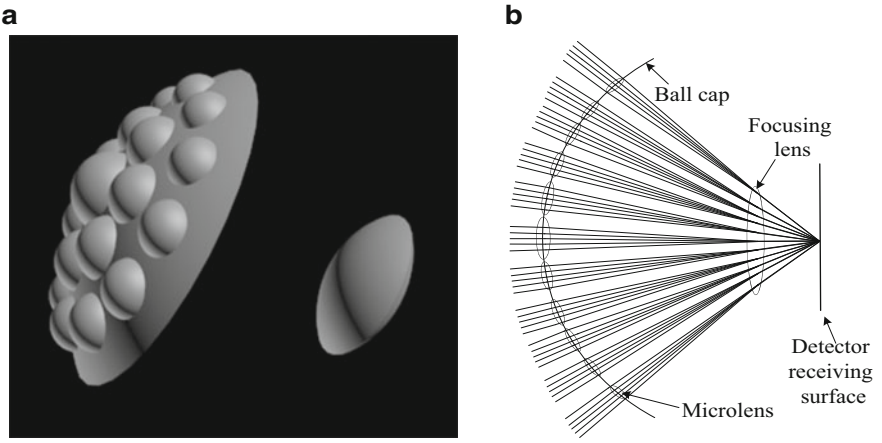


Fig. 31.5 Schematic diagram of a double-curved compound eye lens array optical system. (**a**) Double-layer curved compound eye lens structure. (**b**) Ray tracing of double-curved compound eye lens

To improve the beam convergence effect of the microlens at the edge of the curved structure, further concentrate the beam received by the receiving antenna, and address the defocusing phenomenon in the beam convergence process, a double-layer curved compound lens array optical system was designed based on a single layer.

Figure 31.5 shows a schematic diagram of the optical system of the double-layered curved compound eye lens array. First, the focal length of each small curved lens on the first layer of the system is determined by the position on the semicircle where the lens is located. Thus, small lenses at any position on the surface of the large lens can achieve an ideal beam convergence effect. Second, the light signal can be focused on the surface of the detector once it passes through the large curved lens on the second layer, thereby reducing the light energy loss caused by the inability of the edge lens to focus.

To apply this optical receiving antenna system to the underwater blue-green LED communication system, based on the geometric optical design and simulation of the lens array, the outer spherical crown model of the double-layer compound eye structure was processed, and the small lenses were inlaid and combined with the focusing lens. The actual object is shown in Fig. 31.6.

The light power on the incident plane of the double-layer composite eye lens can be estimated by measuring the power of the LED light source, based on the Lambert radiation model. As shown in Table 31.10, with an increase in the incident angle of the surface light source of the compound eye lens, the output power of the pinhole in front of the power meter decreases. This indicates that the insertion loss of the optical receiving antenna of the double-layer compound eye lens increases with an increase in the incident angle.

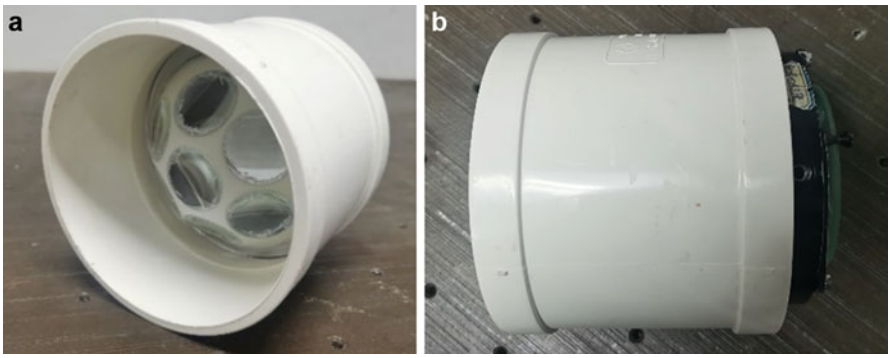


Fig. 31.6 Receiving end prototype with a double-layer compound eye array. (**a**) Front view. (**b**) Vertical view

Table 31.10 Incidence angle and insertion loss of the double-layer curved compound lens

Angle of incidence	0°	5°	10°	15°	20°
Incident power of compound eye lens (W)	4.89×10^{-5}	4.89×10^{-5}	4.89×10^{-5}	4.89×10^{-5}	4.89×10^{-5}
Effective area detector power (W)	3.1×10^{-5}	2.77×10^{-5}	2.17×10^{-5}	1.97×10^{-5}	1.69×10^{-5}
Insertion loss (dB)	1.97	2.48	3.51	3.95	4.60

Figure 31.7 shows the optical power values measured by the optical power meter at different distances and receiving terminals. It can be concluded that when the communication distance is fixed, the optical power received by the double-layer compound eye lens receiving antenna is optimal, followed by that of the single-layer compound eye lens receiving antenna.

When there is no compound eye lens receiving antenna, the optical power measured by the receiving terminal is very weak. When the receiving antenna remains unchanged, the optical power measured at the receiving end decreases with an increase in the communication distance; however, the optical power received by the double-layer compound eye lens receiving antenna is still higher than that of the other two cases.

Figure 31.8 shows the bit error rate of the communication system at different transmission distances. It can be seen from Fig. 31.8 that the bit error rate of the system increases with an increase in the communication distance. The bit error rate was 1.36×10^{-14} when the communication distance was 1 m. When the communication distance reached 3 m, the bit error rate also increased. When the communication distance reached 5 m, the bit error rate increased to 4.58×10^{-12}. In summary, the signal quality attenuation increases with an increase in the transmission distance.

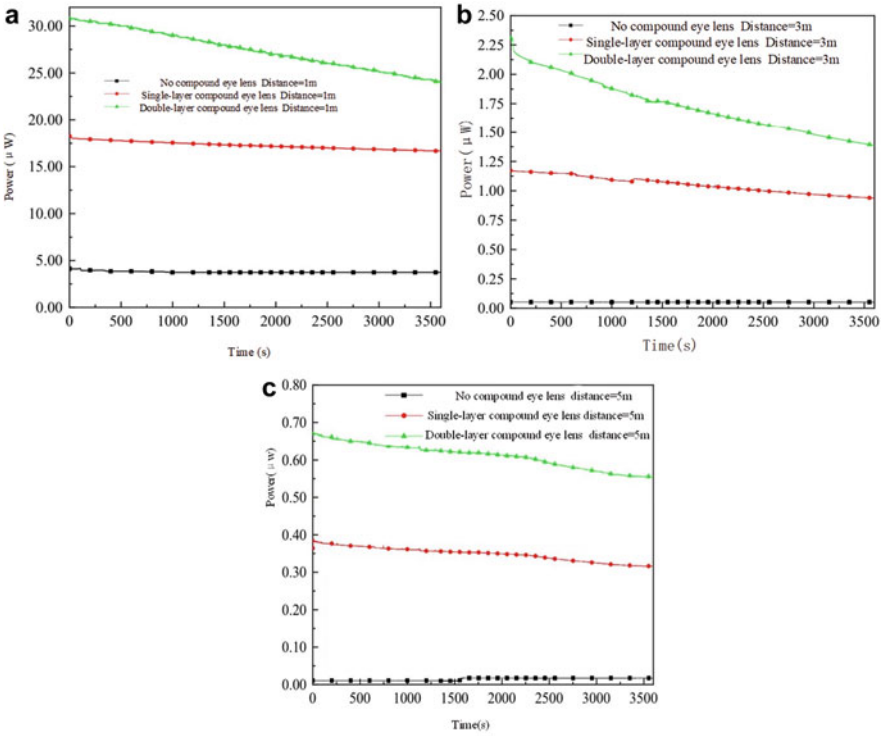


Fig. 31.7 Optical power measurements of different receiving terminals at 1 m, 3 m, and 5 m. (**a**) Power value at 1 m. (**b**) Power value at 3 m. (**c**) Power value at 5 m

Numerical simulation software was used to theoretically calculate the received optical power of the compound eye lens so that by comparing the theoretical data with the experimental data, the difference in optical power received by the receiver with no compound eye receiving antenna, a single-layer compound eye receiving antenna, and a double-layer compound eye receiving antenna can be seen more intuitively.

Because the experimental environment was still water, the influence of turbulence was not considered. The numerical parameters are consistent with the experimental design data: $P_0 = 500$ μW, $D_t = 38$ mm, $D_r = 100$ mm, $\sigma_r^2 = 10^{-10}$–10^{-13}, $\eta_t = 0.82$, $c(\lambda) = 0.151$, $B = 150$ MHz, $T = 290$, $R = 50$ Ω, and $\beta = 20°$. The calculation results are shown in Table 31.11, where P_s is the theoretical data and P_e is the mean value of the experimental data collected through experimental measurements.

Figure 31.9 compares the mean values of the fitted theoretical data and the experimentally measured data, based on the data in Table 31.11. The solid line in Fig. 31.9 is the curve fitted by the experimental data measured continuously by the optical power meter within 1 h. The dotted line is the curve fitted using the formula for the received light power of the double-layer compound eye lens receiving antenna.

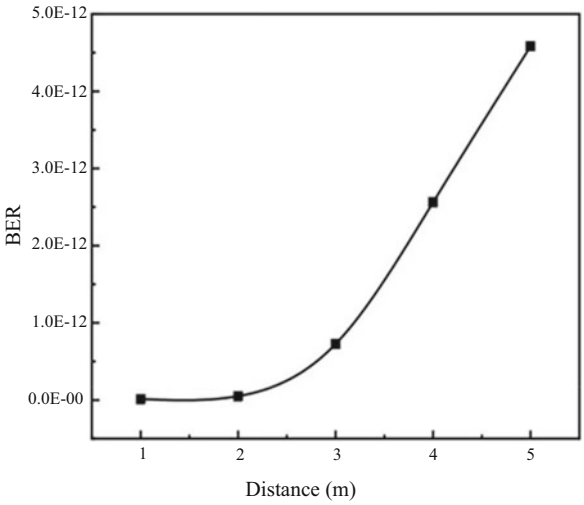


Fig. 31.8 Bit error rate at different communication distances

Table 31.11 Theoretical calculation data and experimental data

Distance	No compound eye lens (W)	Single-layer compound eye lens (W)	Double-layer compound eye lens (W)	Improvement of double-layer compound eye lens over single-layer compound eye lens (%)	Deviation (%)
1 m	$P_s = 4.771$	$P_s = 19.855$	$P_s = 30.762$	72	0.7
	$P_e = 3.765$	$P_e = 19.457$	$P_e = 33.431$	72.7	
3 m	$P_s = 0.048$	$P_s = 1.012$	$P_s = 1.683$	66.3	1.3
	$P_e = 0.052$	$P_e = 1.049$	$P_e = 1.731$	65	
5 m	$P_s = 0.014$	$P_s = 0.303$	$P_s = 0.485$	60.05	0.05
	$P_e = 0.018$	$P_e = 0.378$	$P_e = 0.605$	60	

According to the trend of the solid and dotted lines in Fig. 31.9, these three groups of curves are in line with the exponential decay trend. It can be seen from the figure that when the transmission distance is fixed, the power received by the double-layer compound eye lens is higher than that of the single-layer compound eye lens.

The experimental results show that when the link distances are 1, 3, and 5 m, the received optical powers of the double-layered compound eye system are 72%, 65%, and 60% higher than those of the single-layered system, respectively. A double-layer fly eye lens receiving antenna can significantly improve the optical power received at the receiving end. Therefore, the proposed antenna structure has strong practicability and good development prospects in the field of underwater optical wireless communication.

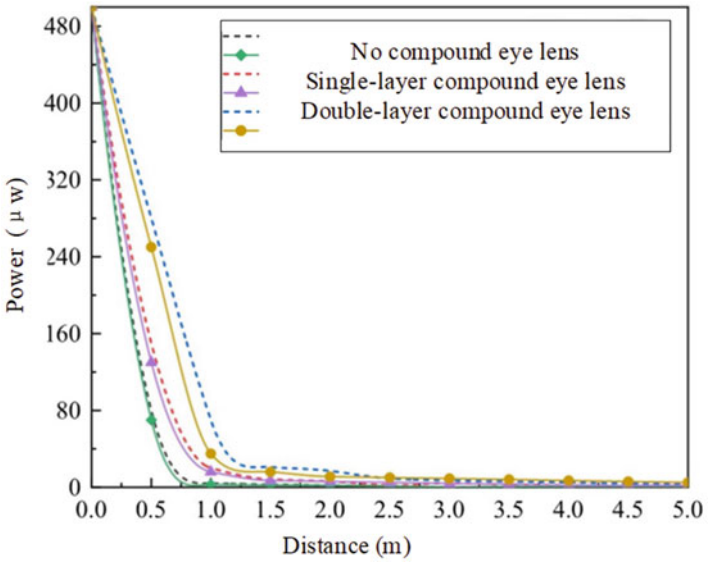


Fig. 31.9 Experimental and theoretical data curves

The double-layered compound eye receiving antenna provides a feasible solution for increasing the underwater optical wireless communication distance. Future research can focus on improving the transmission performance of communication systems by using channel equalization technology.

Summary

Considering the basic theory of underwater optical wireless communication, this chapter analyzed the advantages of underwater optical communication. First, a model for underwater optical communication was introduced. Then, the optical properties of seawater were analyzed, and factors affecting underwater transmission channels were proposed, including absorption, scattering, and turbulence.

Then, the key technologies used in underwater optical communication were introduced, including channel modeling, modulation and coding technology, light source selection, and detection technology. The chapter summarized the research progress and experiments in these aspects and finally introduced a new underwater optical system design model in detail, which provided a feasible solution for underwater long-distance transmission.

According to the research progress in this chapter, many aspects in underwater optical wireless communication systems need to be improved and perfected:

1. The experiments mentioned in this paper are mostly applicable to simplex communication. In the future, high-bandwidth LEDs and high-sensitivity

photodetectors can be combined to achieve half- or full-duplex communication systems.

2. The receiving end of an underwater optical communication system is mostly a single receiving detector. In the future, multiple detectors can be used at the receiving end, and multiple focusing lenses can be added, which can increase the communication distance and reduce the bit error rate [145].
3. An adaptive coding modulation scheme could be designed in the future. In underwater optical wireless communication systems, appropriate modulation and coding methods can be selected according to the actual situation so that the system can save energy and have a longer working time [33].
4. For the cross-medium transmission model, most vertical link experiments assume that seawater is a homogeneous medium; the model can be further studied according to real ocean conditions in the future [183].
5. Owing to the weak optical signal strength received underwater and the fact that the current underwater weak light detection technology is not yet mature, a set of underwater weak light signal detection systems can be studied in the future [212].
6. With the development of the Internet of things and 6G, underwater optical communication is expected to connect with the Internet in the future. Thus, new network protocols will need to be proposed, and underwater acoustic, RF, and optical communication will need to be combined to overcome the problems of unstable channels and low transmission rates.

References

1. Zeng Z, Fu S, Zhang H, et al. A Survey of Underwater Optical Wireless Communications [J]. IEEE Communications Surveys & Tutorials, 2017, 19(1):204–238.
2. Mccartney A. Acoustic Communication, 2nd edition[J]. Canadian Journal of Communication, 2001, 26(03): 459–461.
3. Lan F. Akyildiz, Dario Pompili, Tommaso Melodia. Underwater acoustic sensor networks: research challenges[J]. Ad Hoc Networks, 2005, 3(3): 257–279.
4. Ma Xueli. Research on High Speed Underwater Acoustic Communication Technology [D]. Harbin: Harbin Engineering University, 2004.
5. Kathleen M. Stafford, Christopher G. Fox, Bruce R. Mate. Acoustic detection and location of blue whales (Balaenoptera musculus) from SOSUS data by matched filtering[J]. Acoustical Society of America, 1994,96(5):3250–3251.
6. Chen Dongrong, Tang Yonglu, Chen Xianwen, et al. Convergence of sound field near the track axis [J]. Acoustic Technology, 1987, 3 (03): 15–21.
7. Wang Xiaofeng, Ma Cheng, Qian Zhengfang, et al. Exploration and Suggestions on Calculating Depth in the Design of Large Submersible Submarines [J]. Ship Science and Technology, 2004,7 (03): 13–15.
8. Xiao Peng. Research on sound propagation and sound field characteristics in deep-sea convergence area and shadow area [D]. Xi'an: Northwest Polytechnical University, 2019.
9. Ma Xiaoxiao. Overview of Underwater Communication Technology [J]. Electronic World, 2020, 20 (14): 104–105.
10. Zhang Yinbing. Research and Implementation of Underwater Acoustic Remote Control Safety Detonation System [D]. Xi'an: Northwest University of Technology, 2007.
11. D. B. Kilfoyle, A. B. Baggeroer. The state of the art in underwater acoustic telemetry[J]. IEEE Journal of Oceanic Engineering, 2000,25(1): 4–31.

12. Zhu Yongjian. On the Application of Underwater Acoustic Communication in Submarine Communication [J]. Ship Electronics Engineering, 2003, 6 (03): 51–59.
13. M. Stojanovic. Recent advances in high-speed underwater acoustic communications[J]. IEEE Journal of Oceanic Engineering,1996,21(2):125–136.
14. Zhu S, Chen X, Liu X, et al. Recent Progress in and Perspectives of Underwater Optical Wireless Communication [J]. Progress in Quantum Electronics, 2020, 73(20): 100274–100326.
15. Jiang S. Overview of underwater acoustic communication[J]. IEEE Access, 2018,39(08): 233–244.
16. Au W, Nachtigall P E, Pawloski J L. Acoustic effects of the ATOC signal (75 Hz, 195 dB) on dolphins and whales[J]. Acoustical Society of America Journal, 1997, 101(5):2973–2977.
17. Partan J, Kurose J, Levine B N. A Survey of Practical Issues in Underwater Networks[J]. Acm Sigmobile Mobile Computing & Communications Review, 2007, 11(4):23–33.
18. Chen Sijian. Development and Trend of Overseas Submarine Communication Technology [J]. Military Communication Technology, 1985, 20 (03): 29–38.
19. Li Youcai, Gong Zhengxia, Shi Boyong. Discussion on Air Attack Defense of Long Wave VLF Communication Fixed Stations [J]. Ship Electronic Countermeasure, 2017,40 (01): 39–42.
20. Li Kai. Space borne VLF emission and propagation [J]. Radio Wave and Antenna, 1998,10 (01): 59–62.
21. Zhou Jingbo, Hu Bo, Wei Yuan. Analysis of Air Attack Weapons during the bombing of US military strongholds [J]. Electronic Countermeasures, 2015, 8 (3): 42–47.
22. Liu Weiguo. Application Status of Foreign Modern Ship Communication Technology [J]. Ship Electronics Engineering, 2001, 25 (04): 43–63.
23. Hu Guoqiang, Sheng Jianbing. Basic Application of Submarine Communication Technology [J]. Ship Electronic Countermeasure, 2007, 42 (02): 83–87.
24. Lan Yun, Li Jia, Xing Weisheng. Submarine underwater communication technology [J]. Modern Ships, 2009, 18 (10): 21–27.
25. Fan Lidong. Key Technology of VLF Communication Reconnaissance System [D] Harbin: Harbin Engineering University, 2021.
26. Shi Boyong, Xu Hongqiang. Implementation of VLF communication in complex electromagnetic environment [J] Modern Communication Technology, 2015:16 (2): 1–3.
27. Hu Shumou. Research on the feasibility method of VLF communication in complex electromagnetic environment [J] China Broadband, 2021, 10 (12): 123–128.
28. Liu Chang. USAF Seeks Universal VLF/LF Communication Technology for Strategic Forces [J]. Missile and Space Launch Technology, 2017, 21 (5): 23–28.
29. Jia Xinduo. Common means of submarine communication and its future development [J]. Electronic World, 2012, 66 (13): 28–29.
30. Zhu Hairong, Zhang Bentao, Chen Biao. Propagation Characteristics of Ultra Low Frequency Electromagnetic Waves in Seawater [J]. Command Control and Simulation, 2013, 35 (06): 54–57.
31. Xianhui Che, Lan Wells, Gordon Dickers. Re-Evaluation of RF Electromagnetic Communication in Underwater Sensor Networks[J]. IEEE Communications Magazine, 2010, 48(12): 143–151.
32. Qu Xiaohui. Views on the Development Direction of Domestic Submarine Communication [J]. Ship Electronics Engineering, 2000, 56 (05): 42–44.
33. Liu Pengzhan, Wang Linning, Hu Fangren, et al. Research and prospect of underwater optical communication technology [J]. Digital Ocean and Underwater Attack and Defense, 2022,5 (4): 329–334.
34. Ali M F, Jayakody D N K, Chursin Y A, et al. Recent advances and future directions on underwater wireless communications[J]. Archives of Computational Methods in Engineering, 2020,27(5): 1379–1412.

35. Che X, Wells I, Dickers G, et al. Re-evaluation of RF electromagnetic communication in underwater sensor networks[J]. IEEE Communications Magazine, 2010, 48(12): 143–151.
36. Cochenour B, Dunn K, Laux A, et al. Experimental measurements of the magnitude and phase response of high-frequency modulated light underwater[J]. Applied optics, 2017, 56(14): 4019–4028.
37. Kaushal H, Kaddoum G. Underwater Optical Wireless Communication[J]. IEEE Access, 2016, 4(01):1518–1547.
38. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research Progress of Adaptive Optics in Optical Wireless Communication System of Xi'an University of Technology [J]. Intense Laser and Particle Beam, 2021,33 (08): 30–52.
39. Duntley S Q. Light in the Sea[J]. Journal of the Optical Society of America, 1963, 53(2): 214–233.
40. Gordon Burrows. Remote Sensing of the Environment[J]. The Journal of the Institution of Electronic and Radio Engineers, 1968, 50(25): 531–532.
41. Phillips D M, Abbot R H, Penny M F. Remote Sensing of Sea Water Turbidity with An Airborne Laser System [J]. Phys. D,1971,17(8) :1749–1758.
42. Karp, S. Optical Communications Between Underwater and Above Surface (Satellite) Terminals[J]. IEEE Transactions on Communications, 1976, 24(1):66–81.
43. Doherty J J. Submarine communication[J]. The Journal of the Acoustical Society of America, 1977,69(6): 335–339.
44. Wiener T F, Karp S. The Role of Blue/Green Laser Systems in Strategic Submarine Communications[J]. IEEE Transactions on Communications, 1980, 28(09):1602–1607.
45. Ke Xizheng, Deng Lijun. Optical Wireless Communication [M]. Beijing: Science Press, 2016: 145–148
46. Brian Billard , Peter J. Wilsen. Sea Surface and Depth Detection in the WRELADS Airborne Depth Sounder [J]. Appl. Opt., 1986,25(13):2059–2066.
47. Snow J B, Flatley J P, Freeman D E, et al. Underwater propagation of high-data-rate laser communications pulses[J]. Proceedings of SPIE-The International Society for Optical Engineering, 1992, 1750(12): 419–431.
48. Bales J W, Chrissostomidis C. High bandwidth low-power short-range optical communication underwater[J]. International symposium on Unmanned Untethered Submersible Technology,1995,32(05).406–415.
49. Tivey M, Fucile P, Sichel E. A low power, low cost, underwater optical communication system [J]. Ridge Events, 2000, 1(01): 27–29.
50. Felix Schill, Uwe R. Zimmer, Jochen Trumpf. Visible spectrum optical communication and distance sensing for underwater applications. [J]. Applied Optics, 2003, 16 (01): 46–56.
51. Akyildiz I F, Pompili D, Melodia T. Challenges for efficient communication in underwater acoustic sensor networks[J]. ACM Sigbed Review, 2004, 1(2): 3–8.
52. Farr N, Chave A, Freitag L, et al. Optical modem technology for seafloor observatories [J]. IEEE Access, 2005,62(1):928–934.
53. Cochenour B, Mullen L, Laux A. Phase Coherent Digital Communications for Wireless Optical Links in Turbid Underwater Environments[J]. IEEE Journal of Oceanic Engineering, 2007,9(20):1–5.
54. Hanson F, Radic S. High bandwidth underwater optical communication[J]. Applied Optics, 2008, 47(2):277–283.
55. W. C. Cox, J. A. Simpson, C. P. Domizioli, et al. An underwater optical communication system implementing Reed–Solomon channel coding[J]. Applied Optics,2008,7385(10):1–6.
56. Jaruwatanadilok S. Underwater Optical Wireless Communication Channel Modeling and Performance Evaluation using Vector Radiative Transfer Theory[J]. IEEE Journal on Selected Areas in Communications, 2008, 26(9):1620–1631.
57. Doniec M W, Vasilescu I, Chitre M, et al. AquaOptical: A Lightweight Device for High-rate Long-range Underwater Point-to-Point Communication[J]. IEEE Photonics Journal, 2009, 44(4):1–6.

58. Cochenour B M, Mullen L J, Laux A E. Characterization of the Beam-Spread Function for Underwater Optical Wireless Communication Optical Wireless Communication s Links [J]. IEEE Journal of Oceanic Engineering, 2008, 33(4):513–524.
59. J. A. Simpson, W. C. Cox, J. R. Krier, et al. 5 Mbps optical wireless communication with error correction coding for underwater sensor nodes[J]. IEEE Photonics Technology Letters,2010, 9(23):6–9.
60. W. C. Cox, K. F. Gray, J. A. Simpson, et al. A MEMS Blue/Green Retroreflecting Modulator for Underwater Optical Communications[J]. Optics Express, 2010, 4432(88):1–4.
61. Vucic J, Kottke C, Nerreter S, et al. 513 Mbit/s Visible Light Communications Link Based on DMT-Modulation of a White LED[J]. Journal of Lightwave Technology, 2010, 28(24): 3512–3518.
62. Gabriel C, Khalighi A, Bourennane S, et al. Optical communications systems for an underwater wireless sensor network[J]. The International Society for Optical Engineering, 2012, 7901(01):11–12.
63. Cossu G, Corsini R, Khalid A M, et al. Experimental demonstration of high speed underwater visible light communications[J]. IEEE Journal of Oceanic Engineering, 2013,64(27): 11–15.
64. Doniec M, Angermann M, Rus D. An End-to-End Signal Strength Model for Underwater Optical Communications[J]. IEEE Journal of Oceanic Engineering, 2013, 38(4):743–757.
65. Seongwon Han, Youngtae Noh, R. Liang, et al. Evaluation of underwater optical-acoustic hybrid network[J]. China Communications, 2014,11(5): 49–59.
66. Hiskett P A, Lamb R A. Underwater optical communications with a single photon-counting system[J]. Optics Express, 2014,28(11): 91–106.
67. Oubei H M, Duran J R, Janjua B, et al. 48 Gbit/s 16-QAM-OFDM transmission based on compact 450-nm laser for underwater Optical Wireless Communication [J]. Optics Express, 2015, 23(18):23302–23309.
68. Nakamura K, Mizukoshi I, Hanawa M. Optical wireless transmission of 405 nm, 1.45 Gbit/s optical IM/DD-OFDM signals through a 4.8 m underwater channel[J]. Optics Express, 2015, 23(2):1558–1566.
69. Oubei H M, Li C, Park K H, et al. 2.3 Gbit/s underwater Optical Wireless Communication s using directly modulated 520 nm laser diode[J]. Optics Express, 2016, 23(16):20743–20751.
70. Stephen P. Najda, Piotr Perlin, Tadek Suski, et al. AlGaInN laser diode technology for GHz high-speed visible light communication through plastic optical fiber and water[J]. Optical Engineering. 2016, 55(2):026112–026118.
71. Abdullah Al-Halafi, Hassan Makine Oubei, Boon S. Ooi, et al. Real-Time Video Transmission Over Different Underwater Wireless Optical Channels Using a Directly Modulated 520nm Laser Diode[J]. Journal of Optical Communications and Networking, 2017, 9(10): 826–832.
72. Tsai-Chen Wu, Yu-Chieh Chi, Huai-Yung Wang, et al. Blue Laser Diode Enables Underwater Communication at 12.4 Gbps Scientific Reports[J]. Open Access, 2017, 7(8):40480–40490.
73. Oubei H M, Zedini E, ElAfandy R T, et al. Simple statistical channel model for weak temperature-induced turbulence in underwater Optical Wireless Communication systems [J]. Optics Letters, 2017, 42(13): 2455–2458.
74. Jamali M V, Salehi J A, Akhoundi F. Performance Studies of Underwater Optical Wireless Communication Systems with Spatial Diversity: MIMO Scheme[J]. IEEE Transactions on Communications, 2017, 65(3):1176–1192.
75. A. Al-Halafi, B. Shihada. UHD Video Transmission Over Bidirectional Underwater Optical Wireless Communication [J]. IEEE Photonics Journal, 2018, 10(2):1–14.
76. Illi E, Bouanani F E, Ayoub F. A high accuracy solver for RTE in underwater optical communication path loss prediction[J]. Optical Engineering, 2018, 23(9):1–8.
77. Illi E, Bouanani F E, Park K H, et al. An Improved Accurate Solver for the Time-Dependent RTE in Underwater Optical Wireless Communications[J]. IEEE Access, 2019, 68(99):1–3
78. Tsai W S, Lu H H, Wu H W, et al. A 30 Gb/s PAM4 underwater wireless laser transmission system with optical beam reducer/expander[J]. Scientific Reports, 2019, 9(01):8605–8613.

79. Ramavath P N, Udupi S A, Krishnan P. High-speed and reliable Underwater Optical Wireless Communication system using Multiple-Input Multiple-Output and channel coding techniques for IoUT applications[J]. Optics Communications, 2020, 461(10): 125229–125229.
80. Arvanitakis G N, Bian R, Mckendry J, et al. Gb/s Underwater Optical Wireless Communication s Using Series-Connected GaN Micro-LED Arrays[J]. IEEE Photonics Journal, 2020, 12(2):1–10.
81. Kong M, Guo Y, Alkhazragi O, et al. Real-time Optical-Wireless Video Surveillance System for High Visual-fidelity Underwater Monitoring[J]. IEEE Photonics Journal, 2022, 14(2): 1–11.
82. Jiang Liu, Tan Xinquan, Yao Tianren, et al. Image signal processing method in laser underwater communication [J]. Journal of Huazhong University of Science and Technology, 1987, 25 (12): 23–29.
83. Liang Chunyan, Zou Chuanyun, Ao Faliang. Implementation of RS code frequency domain error correction coding in laser submarine communication channel [J]. Telecommunication Technology, 1989,38 (06): 45–49.
84. Sui Meihong, Yu Xinsheng, Liu Xifeng, et al. Research on characteristics of underwater optical wireless communication channel [J]. Ocean Science, 1992, 33 (06): 80–85.
85. Wang Fuchang, Huang Benxiong, He Dake. Laser underwater PPM communication slot synchronizer [J]. Journal of Huazhong University of Technology, 1995, 23 (10): 52–56.
86. Yang Min, Huang Benxiong, Wang Fuchang, et al. Application of error correction code in laser underwater communication [J]. Journal of Huazhong University of Science and Technology, 1997, 25 (07): 11–13.
87. Huang Xiaosheng, Wang Rulin, Xu Rensheng, et al. Underwater laser communication transmission and reception system [J]. Journal of Qingdao Ocean University (Natural Science Edition), 1998116 (04): 140–145.
88. Zou Chuanyun, Ao Faliang, Huang Xiangfu. Adaptive enhancement of diversity multiplex signal in underwater laser communication [J]. Telecommunication Technology, 1999,42 (04): 36–40.
89. He Ning, Li Hailing, Zhang Dekun, et al Diversity reception of signals in underwater laser communication [J] Laser and Infrared, 2002, 32 (04): 228–229.
90. Wang Kun, Zhang Dekun. Application of fiber grating filter in underwater optical communication system [J]. Optical Communication Research, 2003 (03): 32–34.
91. Pang Hanhan. Design and Implementation of High Speed Underwater Optical Communication System Based on LED [D]. Qingdao: Ocean University of China, 2004.
92. Zhou Tianhua, He Ning, Ao Faliang. DSP Implementation of PPM Digital Receiver for Underwater Optical Communication [J]. Optical Technology, 2006, 32 (51): 607–609.
93. Cui Zhun. Research on Polarization Technology in Underwater Laser Communication [J]. Telemetry and Remote Control, 2008, 29 (06): 63–67.
94. Luo Linfeng, Wang Shuo. Design Scheme of Optical Communication and Ranging in Underwater Mobile Sensor Network [J] Robot Technology and Application, 2008, 02 (02): 36–39.
95. He Yan, Shang Jianhua, Liu Dan, et al. Heterodyne laser Doppler vibrometer for underwater acousto-optic communication [J]. China Laser, 2009, 36 (01): 189–192.
96. Liu Jintao, Chen Weibao. Research on satellite uplink laser communication by underwater platform [J]. Acta Photonica Sinica, 2010, 39 (04): 693–698.
97. Cong Yanping. Research on key technologies of multi-mode adaptive underwater wireless communication network [D]. Qingdao: Ocean University of China, 2011.
98. Wang Feng, Rao Jionghui. Realization of variable gain amplifier in underwater Optical Wireless Communication system [J]. Laser Technology, 2012,36 (01): 99–102.
99. Ao Jun, Zhu Yong, Wang Zheng, et al. Underwater PTZ control system based on electronic compass [J]. Naval Electronics Engineering, 2012, 32 (04): 79–82.
100. Tian B, Zhang F, Tan X. Design and development of an LED-based optical communication system for autonomous underwater robots[J]. IEEE Access, 2013, 7(12) 1558–1563.

101. Xu J, Kong M, Lin A, et al. OFDM-based broadband underwater Optical Wireless Communication system using a compact blue LED[J]. Optics Communications, 2013, 369(01): 100–105.
102. Hu Xiuhan, Zhou Tianhua, He Yan, et al. Design and analysis of underwater optical communication transceiver system based on digital signal processor [J]. China Laser, 2013, 40 (03): 128–134.
103. Xi H, Yan J. Study on performance of M-ary PPM underwater optical communication systems using vector radiative transfer theory[J]. IEEE Photonics Journal, 2014, 22(26): 566–570.
104. Cai Wenyu, Wen Duanqiang, Fang Xun. Research and Implementation of Underwater LED Optical Communication Technology [J]. Optical Communication Technology, 2015, 39 (06): 60–62.
105. Shen C, Guo Y, Oubei H M, et al. 20-meter underwater Optical Wireless Communication link with 1.5 Gbps data rate[J]. Optics Express, 2016, 24(22):25502–25510.
106. Hu Feng, Wu Guofeng, Wu Gang, et al. Design of underwater two-way communication test system based on blue LED [J]. Electronic Design Engineering, 2016, 24 (08): 123–130.
107. Du Jinsong, Zhou Tianhua, Chen Weibao, et al. Performance Analysis of Underwater Optical Communication Based on LDPC and PPM [J]. Progress in Laser and Optoelectronics, 2016, 53 (12): 100–106.
108. H.-H. Lu, C. Y. Li, H. H. Lin, et al. An 8m/9.6Gbps underwater Optical Wireless Communication system[J], IEEE Photonics Journal, 2016, 8(5):1109–1119.
109. Jing Xu, Yuhang Song, Xiangyu Yu, et al. Underwater wireless transmission of high-speed QAM-OFDM signals using a compact red-light laser[J]. Optics Express,2016,24(8): 8097–8109.
110. Liu X, Yi S, Zhou X, et al. 34.5 m underwater optical wireless communication with 2.70 Gbps data rate based on a green laser diode with NRZ-OOK modulation[J]. Optics Express, 2017, 25(22):27937–27947.
111. Chen Y, Kong M, Ali T, et al. 26 m/5.5 Gbps air-water optical wireless communication based on an OFDM-modulated 520-nm laser diode[J]. Optics Express, 2017, 25(13):14760–14765.
112. Ao Jun, Yao Meng, Ma Chunbo. Design of underwater LED optical communication system based on MIMO [J]. Optical Communication Technology, 2017, 41 (09): 24–31.
113. Lin Aobo. FPGA based underwater OFDM wireless visible light communication system [D]. Hangzhou: Zhejiang University, 2019.
114. Song Yuhang. Underwater wireless visible light communication system based on MIMO-OFDM [D]. Hangzhou: Zhejiang University, 2017 [115] Huang Y F, Tsai C T, Chi Y C, et al. Filtered Multicarrier OFDM Encoding on Blue Laser Diode for 14.8-Gbps Seawater Transmission[J]. Journal of Lightwave Technology, 2018, 65(99):1–7.
115. Wang Hongxi, He Fengtao, Zhan Fei, et al. Design of underwater LED blue light communication system based on PPM [J]. Optical Communication Technology, 2018, 42 (07): 46–50.
116. Han Biao. Research on underwater two-way Optical Wireless Communication for deep-sea applications [D]. Xi'an: University of Chinese Academy of Sciences (Xi'an Institute of Optics and Precision Machinery, Chinese Academy of Sciences), 2018.
117. Yang Jie, Wang Jianfeng, Chen Qi. Design and simulation of an underwater optical communication system based on PPM modulation [J]. Computer and Digital Engineering, 2018, 46 (07): 1473–1477.
118. P. Wang, C. Li, Z. Xu. A Cost-Efficient Real-Time 25 Mb/s System for LED-UOWC: Design, Channel Coding, FPGA Implementation, and Characterization[J]. Journal of Lightwave Technology, 2018,36(13):2627–2637.
119. Zhang Liangliang. Research on multi-channel underwater optical communication system based on PWM modulation method [D]. Harbin: Harbin University of Technology, 2018.
120. Li Yanfang. Research on underwater high-speed wireless optical two-way communication system [D]. Dalian: Dalian University of Technology, 2018.

121. Jiemei Wang, Chunhui Lu, Shangbin Li, et al. 100 m/500 Mbps underwater optical wireless communication using an NRZ-OOK modulated 520 nm laser diode [J]. Optics Express, 2019, 27(09): 12171–12181.
122. Hong X, Fei C, Zhang G, et al. Discrete multitone transmission for underwater optical wireless communication system using probabilistic constellation shaping to approach channel capacity limit[J]. Optics Letters, 2019, 44(3):558–561.
123. Jiang Xiao, Fu Yu, Kang Linchong, etc Current Situation and Prospect of Ocean IoT Technology [J] Ocean Information, 2019, 34 (3): 7–11.
124. J. Wang, C. Tian, X. Yang, et al. Underwater Optical Wireless Communication system using a 16-QAM modulated 450-nm laser diode based on an FPGA[J]. Applied Optics, 2019,58(16): 4553–4559.
125. Y. Zhou, X. Zhu, F. Hu, et al. Common-anode LED on a Si substrate for beyond 15 Gbit/s underwater visible light communication[J]. Photonics Research, 2019,7(9):1019–1033.
126. Meng Linhu. Research on Underwater Laser Communication Technology [D]. Xi'an: Xi'an University of Technology, 2021.
127. Han Mengxiang. Experimental research on high-speed underwater wireless communication based on blue-green laser [D]. Beijing: Beijing University of Posts and Telecommunications, 2021.
128. Shang Yawen. Research on underwater real-time wireless communication technology based on blue-green laser [D]. Beijing: Beijing University of Posts and Telecommunications, 2021.
129. J. Li, B. Yang, L. Wang, et al. A Real-Time, Full-Duplex System for Underwater Optical Wireless Communication: Hardware Structure and Optical Link Model[J]. IEEE Access,2020,8(24):109372–109387.
130. Ji D U, Wang Y, Fei C, et al. Experimental demonstration of 50-m/5-Gbps underwater optical wireless communication with low-complexity chaotic encryption[J]. Optics Express, 2020, 29(2): 783–796.
131. Ao Jun, Wu Tong, Ma Chunbo. An underwater distorted image restoration method based on depth learning [J]. Computer Simulation, 2020, 37 (08): 214–218.
132. Kaipeng Li, Yan He, Jian Ma, et al. A Dual-Wavelength Ocean Lidar for Vertical Profiling of Oceanic Backscatter and Attenuation[J]. Remote Sensing, 2020, 12(17):2844–2864.
133. Wu Yixin. Design of underwater positioning and communication integrated system based on photoelectric array [D]. Guilin: Guilin University of Electronic Science and Technology, 2023.
134. Tong Junjie. Underwater high-speed Optical Wireless Communication system based on OFDM [D]. Guilin: Guilin University of Electronic Science and Technology, 2023.
135. Dong Bing, Tong Shoufeng, Zhang Peng, et al Prototype design of 20 m underwater wireless blue LED communication system [J] China Optics, 2021, 14 (06): 1451–1458.
136. Lv Weichao, Xu Jing. Underwater Optical Wireless Communication system based on spread spectrum technology [J]. Radio Communication Technology, 2021,47 (04): 410–415.
137. He Fengtao, Wei Pengcheng, Zhang Jianlei, et al. Design and implementation of underwater wireless optical automatic alignment system [J]. Laser Journal, 2021,42 (01): 124–131.
138. Ma Chunbo, Gu Jinbo, Huang Xinyi, et al. Design of underwater visible video transmission system based on blue LED [J]. Optical Communication Technology, 2021,45 (12): 21–28.
139. Yan Weilong, Li Xiaoji, Wang Junhua, et al. Pre equalization technology based on DCO-OFDM underwater visible light communication system [J]. Optical Communication Technology, 2021,45 (04): 12–15.
140. Li Jinjia, Ye Demao, Wang Linning, et al. Application of PMT array in underwater MIMO Optical Wireless Communication [J]. Infrared and Laser Engineering, 2021,50 (08): 300–308.
141. Chai Chenyang. Design and implementation of underwater visible light OFDM communication system [D]. Guilin: Guilin University of Electronic Science and Technology, 2024.
142. Shen Tianhao. Research on underwater wireless communication system based on blue-green visible light [D]. Huangshi: Hubei Normal University, 2024.
143. C. Fei, Y. Wang, J. Du, et al. 100-m/3-Gbps underwater wireless optical transmission using a wideband photomultiplier tube (PMT)[J]. Optics Express, 2022, 30(2): 2326–2337.

144. Ke X Z, Yang S J, Sun Y, et al. Underwater blue-green LED communication using a double-layered, curved compound-eye optical system[J]. Optics Express,2022,30(11):18599–18618.
145. Shi Jia, Huang Aiping, Tao Linwei. Deep Learning Assisted Underwater Optical Communication Channel Estimation and Signal Detection [J]. China Laser, 2022,60 (23): 18–33.
146. Li Bili, He Fengtao, Zhu Yunzhou, et al. Research on high-power underwater laser communication and emission system [J]. Automation and Instrumentation, 2022,8 (02): 30–47.
147. Zhang Jianlei, Lan Xiang, Yang Yi, et al. Adaptive stochastic resonance detection of weak light signal in underwater laser communication [J]. Optical Precision Engineering, 2022,30 (12): 1383–1393.
148. Wang Zhenya. Research on underwater blue-green optical communication system based on APD [D]. Wuhan: Wuhan Textile University, 2026.
149. Chen Wenjiang, Chen Qi'an, Chen Hongming. A Solution for Underwater Optical Wireless Communication System [J]. China Integrated Circuit, 2022,31 (05): 61–66.
150. Smith R C, Baker K S. Optical properties of the clearest natural waters (200–800 nm) [J]. Applied Optics, 1981, 20(2):77–84.
151. Mobley C D, Gentili B, Gordon H R, et al. Comparison of numerical models for computing underwater light fields[J]. Applied Optics, 1993, 32(36):7484–7504.
152. Chao W, Yu H Y, Zhu Y J. A Long Distance Underwater Visible Light Communication System with Single Photon Avalanche Diode[J]. IEEE Photonics Journal, 2017, 8(5):1–11.
153. Yj A, Gw A, Yu Z A. Performance Analysis of SPAD-Based Underwater Optical Wireless Communication Systems[J]. Procedia Computer Science, 2018, 131(04):1134–1141.
154. W. Lu, L. Liu, J. Sun, Influence of temperature and salinity fluctuations on propagation behaviour of partially coherent beams in oceanic turbulence[J]. Applied Optics, 2006, 8(12): 1052–1058.
155. Kolmogorov A N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers[J]. Mathematical and Physical Sciences, 1991, 434(1890): 9–13.
156. Yang Tianxing. Ocean turbulence phase screen model and OAM beam propagation characteristics under this model [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2018.
157. Nikishov V V, Nikishov V I. Spectrum of turbulent fluctuations of the sea-water refraction index[J]. International journal of fluid mechanics research, 2000, 27(1):82–98.
158. Jiang Hongyan. Research on key performance of underwater Optical Wireless Communication in turbulent channel [D]. Guilin: Guilin University of Electronic Science and Technology, 2024.
159. Lu Lu. The influence of ocean turbulence on laser beam transmission [D]. Hefei: University of Science and Technology of China, 2018.
160. Guo Yilin. Study on propagation characteristics of vortex beams in marine turbulent channels [D]. Beijing: Beijing University of Posts and Telecommunications, 2021.
161. Zhang Ting. Study on transmission performance of underwater Optical Wireless Communication system considering the influence of ocean turbulence [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2024.
162. Hamza, Gerekciolu. Bit error rate of focused Gaussian beams in weak oceanic turbulence [J]. Journal of the Optical Society of America A, 2014, 31(9):1963–1968.
163. Tang S, Dong Y, Zhang X. Impulse Response Modeling for Underwater Optical Wireless Communication Links[J]. IEEE Transactions on Communications, 2014, 62(1):226–234.
164. Du Ying. Research on the Performance of PPM Optical Wireless Communication in Ocean Anisotropic Turbulence [D]. Xi'an: Xi'an University of Posts and Telecommunications, 2023.
165. He Fengtao, Li Jiaqi, Zhang Jianlei, et al. Performance Analysis of Wavelength Diversity Optical Wireless Communication System in Ocean Turbulence [J]. Infrared and Laser Engineering, 2021,50 (12): 1–10.
166. Li Y, Zhang Y, Zhu Y, et al. Effects of anisotropic turbulence on average polarizability of Gaussian Schell-model quantized beams through ocean link[J]. Appl Opt, 2016, 55(19): 5234–5239.

167. Zou Z, Wang P, Chen W, et al. Average capacity of a UOWC system with partially coherent Gaussian beams propagating in weak oceanic turbulence[J]. Journal of the Optical Society of America A, 2019, 36(9):1463–1474.
168. Liu L, Zhou S, Cui J H. Prospects and problems of wireless communication for underwater sensor networks[J]. Wireless Communications & Mobile Computing, 2010, 8(8):977–994.
169. Smart J H. Underwater optical communications systems part 1: variability of water optical parameters[J]. IEEE Access, 2005, 2(20):1140–1146.
170. Li C, Park K H, Alouini M S. On the Use of a Direct Radiative Transfer Equation Solver for Path Loss Calculation in Underwater Optical Wireless Channels[J]. IEEE Wireless Communications Letters, 2015, 4(5):561–564.
171. Li J, Ma Y, Zhou Q, et al. Channel capacity study of underwater Optical Wireless Communication s links based on Monte Carlo simulation[J]. Journal of Optics, 2012, 14(1):402–407.
172. Gabriel C, Khalighi M A, Bourennane S, et al. Monte-Carlo-Based Channel Characterization for Underwater Optical Communication Systems[J]. Journal of Optical Communications and Networking, 2012,5(1):1–12.
173. Dalgleish F R, Hou W, Arnone R A, et al. Efficient laser pulse dispersion codes for turbid undersea imaging and communications applications[J]. IEEE Photonics Journal, 2010, 10(25): 7678–7689.
174. Ding H, Xu Z, Sadler B. A Path Loss Model for Non-Line-of-Sight Ultraviolet Multiple Scattering Channels[J]. EURASIP Journal on Wireless Communications and Networking, 2010, 10(2):1–12.
175. Yuan R, Ma J, Su P, et al. Monte-Carlo Integration Models for Multiple Scattering Based Optical Wireless Communication[J]. IEEE Transactions on Communications, 2020, 68(1): 334–348.
176. Arnon S, Kedar D. Non-line-of-sight underwater optical wireless communication network [J]. Journal of the Optical Society, 2009, 26(3):530–539.
177. Tang S, Dong Y, Zhang X. On path loss of NLOS underwater Optical Wireless Communication links[J]. Optics Express, 2013, 6(14):1–3.
178. Jagadeesh V K, Choudhary A, Bui F M, et al. Characterization of Channel Impulse Responses for NLOS Underwater Optical Wireless Communication [J]. Journal of Optics, 2014, 27(29): 77–79.
179. Anous N, Abdallah M, Uysal M, et al. Performance Evaluation of LOS and NLOS Vertical Inhomogeneous Links in Underwater Visible Light Communications[J]. IEEE Access, 2018, 6(09):22408–22423.
180. Zhai Deyan. Study on the transmission characteristics of transmedium laser [D]. Guilin: Guilin University of Electronic Science and Technology, 2024.
181. An Dongdong. Simulation study on capsizing of ships under random wind and waves [D] Harbin: Harbin Engineering University, 2021.
182. Liu Li. Research on cross medium blue green laser uplink communication system [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2021.
183. Li Cong, Yuan Ruiman, Gao Hong, et al. Transmission characteristics of blue-green laser downlink under different weather conditions [J]. Acta Photonica Sinica, 2021, 50 (12): 42–56.
184. Si Lihong. Research on characteristics of blue green laser uplink communication channel [D]. Guilin: Guilin University of Electronic Science and Technology, 2007.
185. Chen M F, Fung A K. A numerical study of the regions of validity of the Kirchhoff and small-perturbation rough surface scattering models[J]. Radio Science, 1988, 23(2):163–170.
186. Wang Min, Liu Weiting, Wang Jianhua. Analysis of signal energy transfer in blue-green laser uplink transmission channel [J] Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2003, 17 (04): 38–41.
187. Li Yinlong, Mao Zhongyang, Liu Xiguo. Analysis of the characteristics of the blue green laser communication channel in the air and underwater [J]. Naval Electronics Engineering, 2021, 41 (09): 69–73.

188. Zhang Jialin, Wang Mingjun, Liu Yongzhi, et al. Scattering characteristics on the blue green laser downlink transmission link of the gas sea dynamic cross medium sea surface [J]. Infrared and Laser Engineering, 2022, 51 (06): 25–31.
189. Wu Chengzhi. Preliminary Exploration of Underwater Optical Wireless Communication System [J]. Modern Transmission, 2014, 15 (05): 72–79.
190. Duan Chengli, Wang Zhen. Research progress of superluminescent diodes [J]. Semiconductor Optoelectronics, 2013, 3 (03): 361–361.
191. Chi Nan, Chen Hui Frontier research progress of high-speed visible light communication [J]. Optoelectronic Engineering, 2020, 47 (03): 6–19.
192. Tian P, Liu X, Yi S, et al. High-speed underwater optical wireless communication using a blue GaN-based micro-LED[J]. Optics Express, 2017, 25(2):1193–1201.
193. Liu X, Yi S, Zhou X, et al. Laser-based white-light source for high-speed underwater Optical Wireless Communication and high-efficiency underwater solid-state lighting[J]. Optics Express, 2018, 26(15):19259–19274.
194. Chao, Fei, Xiaojian, et al. 18.6 Gbps data rate for underwater wireless optical transmission with single laser diode achieved with discrete multi-tone and post nonlinear equalization [J]. Optics Express, 2018, 26(26):34060–34069.
195. Chen X, Lyu W, Zhang Z, et al. 56-m/3.31-Gbps underwater Optical Wireless Communication employing Nyquist single carrier frequency domain equalization with noise prediction [J]. Optics express, 2020, 28(16):23784–23795.
196. Dong Y, Zhang T, Zhang X. Polarized Pulse Position Modulation for Optical Wireless Communication s[J]. Optics Express, 2013, 20(22): 1–5.
197. Huang Y F, Tsai C T, Chi Y C, et al. Filtered Multicarrier OFDM Encoding on Blue Laser Diode for 14.8-Gbps Seawater Transmission[J]. Journal of Lightwave Technology, 2017, 39(99):1–4.
198. Zhang X, Dong Y, Tang S. Polarization differential pulse position modulation[J]. IEEE Photonics Journal, 2012, 5(6): 1–2.
199. Cox W C, Hughes B L, Muth J F. A polarization shift-keying system for underwater optical communications[J]. IEEE Journal of Oceanic Engineering, 2009, 86(2):1–4.
200. Sui M, Yu X, Zhang F. The Evaluation of Modulation Techniques for Underwater Optical Wireless Communication [J]. IEEE Journal of Oceanic Engineering, 2009, 27(28): 138–142.
201. Gabriel C, Khalighi M A, Bourennane S, et al. Investigation of suitable modulation techniques for underwater Optical Wireless Communication [J]. Optics Express, 2012, 60(8):1–3.
202. Ke Xizheng, Yin Zhiyun. Coding Theory in Wireless Laser Communication System [M]. Beijing: Science Press, 2009: 116–187.
203. Liang Jingyuan, Li Mengru, Wang Jiafan, et al. Research progress of error correction coding in Optical Wireless Communication systems [J] Journal of the Internet of Things, 2022, 6 (03): 23–36.
204. Cox W C, Simpson J A, Domizioli C P, et al. An underwater optical communication system implementing Reed-Solomon channel coding[J]. IEEE Photonics Journal, 2008, 9(18): 1–6.
205. Doniec M, Xu A, Rus D. Robust real-time underwater digital video streaming using optical communication[J]. IEEE Access, 2013, 25(10): 5117–5128.
206. Khalighi M A, Uysal M. Survey on free-space optical communication: a communication theory perspective[J]. IEEE Communications Surveys & Tutorials, 2014, 16(4): 2231–2258.
207. Khalighi M A, Akhouayri H, Hranilovic S. Silicon-Photomultiplier-Based Underwater Optical Wireless Communication Using Pulse-Amplitude Modulation[J]. IEEE Journal of Oceanic Engineering, 2019, 45(04):1611–1624.
208. C Fei, J Zhang, G Zhang, et al. Demonstration of 15-M 7.33-Gb/s 450-nm Underwater Wireless Optical Discrete Multitone Transmission Using Post Nonlinear Equalization [J]. Journal of Lightwave Technology, 2018, 36(03): 728–734.
209. Sun M, Zheng B, Zhao L, et al. A design of the video transmission based on the underwater laser communication[J]. IEEE Access, 2014, 9(19):1–4.

210. Li Y, Safari M, Henderson R, et al. Optical OFDM With Single-Photon Avalanche Diode [J]. IEEE Photonics Technology Letters, 2015, 27(9):943–946.
211. Chu Xinyi, Yuan Renzhi, Peng Mugen. Key Technologies and Future Outlook of Underwater Optical Wireless Communication [J]. Mobile Communication, 2022, 46 (06): 86–90.
212. Wang Linning, Liu Pengzhan, Hu Fangren, et al. Underwater optical communication system for 6G network [J]. Mobile Communication, 2022, 46 (06): 45–51.

Experimental Study on Underwater Optical Wireless Channel Models

32

Contents

Introduction 1214
Foreign and Domestic Research Status 1215
Foreign Research Status 1216
Domestic Research Status 1217
Characteristics of Seawater Channels 1219
Ocean Turbulence Spectrum Models 1220
Attenuation Characteristics of Underwater Particles 1221
Experimental Measurements of Underwater Optical Wireless Communication 1232
Experimental System 1232
Statistical Analysis of Data 1234
Underwater Structure Constant 1235
Probability Density Function 1238
Performance Analysis of Underwater Optical Wireless Communication Experiments 1241
Scintillation Index 1241
Optical Power Loss 1242
BER at Different SNR 1244
Optical Cross-Medium Transmission Characteristics 1245
Optical Transmedia Model 1246
Experiment 1248
References 1251

Abstract

When a laser beam propagates underwater, it attenuates because of absorption, scattering, and turbulence; this results in light power attenuation and light intensity fluctuation. Considering that communication links have different water speeds, directions, and distances, an underwater optical wireless channel model is proposed to conduct field measurements in river and ocean nearshore waters. The light intensity fluctuation, received light power loss, and underwater structure constant were measured and analyzed in different water areas. The results showed that rivers have weak turbulence and offshore waters have moderate turbulence. The higher the water speed, the greater the scintillation index of the received

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_32

optical signal power and the more unstable the signal. When transmitting horizontally, the scintillation index of the received optical signal power is independent of the water direction; the greater the scintillation index of the received optical signal power, the more unstable the signal.

Keywords

Underwater optical wireless communication · Channel model · Bit error rate

Introduction

With the deployment of large numbers of unmanned underwater vehicles, the denseness of underwater sensor networks, and the sharp increase in underwater sensor nodes, the demand for underwater wireless data transmissions will increase greatly. Compared with traditional acoustics, optical wireless communication is becoming the main solution for high throughput and big data underwater wireless communication.

However, the serious attenuation loss of underwater optical wireless communication (UOWC) channels, including signal attenuation caused by absorption, scattering, and turbulence, inhibits the transmission of signals. To make UOWC available for a longer link range, it is necessary to conduct a comprehensive study of underwater optical communication and accurate channel modeling.

In 1963, Duntley found a blue-green band with a low-attenuation window (450–550 nm) for the absorption and scattering effects of seawater [1]. In 1966, Gilbert et al. [2] experimentally confirmed this conclusion. In 1972, Petzold divided seawater into four types, according to turbidity, and provided the absorption and scattering coefficients for each type of water [3].

In 2000, Nikishov deduced the power spectrum of the ocean turbulence refractive index, based on the theory that it was affected by both temperature and salinity [4]. In 2008, Jaruwatanadilok modeled an underwater optical wireless channel with absorption and scattering effects and obtained an analytical solution of the radiated transport equation using the vector radiation theory [5]. In the same year, Jaruwatanadilok based on the small-angle approximation hypothesis, a beam diffusion function model was obtained, and the results were verified experimentally [6]. In 2012, Cox used the particle Monte Carlo method to study channels of different water qualities and obtained an optical power attenuation model and a channel impulse response model for different water qualities [7]. In 2013, Li proposed an accelerated algorithm to improve the above method, which reduced the computational complexity [8].

For the first time, Lu et al. derived the wave structure function of plane and spherical waves in ocean turbulence propagation [9]. In 2014, Huang investigated the propagation properties of an underwater Gaussian–Schell model vortex beam under water [10].

In the same year, Lu et al. studied a Gaussian array beam and obtained a distribution function for the average intensity of partially coherent and incoherent light [11]. McGlamery [12] and Roddier [13] proposed power spectrum inversion and Zernike polynomial methods, respectively, to generate a turbulent random-phase screen that satisfied these conditions [14]. In 2021, Cui et al. studied the intensity distribution and scintillation of Gaussian convergent array beams passing through a phase screen. In 2021, Zhang et al. simulated the coexistence of absorption scattering and turbulence using a ray tracing-based particle Monte Carlo numerical method [15].

Many statistical models exist for the random fluctuation of light intensity caused by turbulence, including the log-normal distribution model based on weak turbulence [16] and the gamma–gamma distribution model [17]. In 2017, Oubei proposed a Weibull distribution model [18], and in the same year, Jamali et al. proposed an exponential log distribution model to fit the effect of different bubbles on light intensity fluctuations [19]. In 2019, Zedini et al. developed a mixed exponential generalized gamma function distribution model that considered the effects of temperature, salinity, and bubbles [20].

The above research on laser propagation in seawater is generally based on theoretical analyses and numerical simulations and lacks simulations of real environments. In this study, an underwater channel model was established, based on detailed environmental measurements of nearshore and turbid water bodies. The received light intensity at different flow velocities and flow directions was measured in different turbid waters of the Chan and Ba rivers in Xi'an, Shaanxi Province. In addition, measurements were carried out at different distances in the Yantai No. 1 beach. Compared with the indoor simulation of an underwater environment, the model has more practical significance and reference value and can be used as a reference for future underwater communication work.

Foreign and Domestic Research Status

During the Shang and Zhou dynasties, circa 1600–256 BC, soldiers used beacons to relay information about enemy attacks. Around the first and second centuries, the Roman army continued the practice of Greek soldiers, sending signals using polished shields to reflect sunlight. In 1880, Bell, the father of the telephone, invented a wireless telephone system that used sunlight as a transmission medium. It was the earliest optoelectronic communication system in history.

The advent of the laser in 1916 introduced a large number of optical wireless communication applications to the world. However, because the serious attenuation of visible light in seawater could not be solved, early research on UOWC did not make substantial progress.

With the development of ocean exploration, real-time and high-speed underwater communication has become increasingly important. Underwater optical wireless communication has many advantages, such as high speed, high bandwidth, and large capacity; therefore, an increasing number of researchers have begun to study

it. After more than half a century of development, researchers worldwide have achieved remarkable progress in underwater optical wireless communication.

Foreign Research Status

In 1963, Duntley first proposed a variation in the attenuation coefficient of light waves propagating in different water conditions (pure water, distilled water, and natural water) and found that the attenuation of light in water was weakest when it was in the blue-green light band [21]. This indicated that blue-green light was an important window for light propagation in water, which laid the foundation for the future development of UOWC.

In 1977, Doherty et al. implemented a one-way underwater optical wireless communication system from the shore to a submarine using a focused blue-green laser diode (LD) as the light source [22]. This proved that it was feasible to establish an underwater optical wireless communication system with an LD transmitter to communicate at greater depths and higher cruise speeds.

Between 1980 and 1990, the United States Navy and former Soviet Navy conducted several submarine-to-submarine and satellite-to-submarine underwater optical wireless communication system test studies [23]. They proved that underwater optical wireless communication systems worked well under severe conditions, such as heavy rain and turbid seawater. Since the beginning of this century, the United States, Italy, Japan, and other countries have conducted further research in the field of underwater optical wireless communication.

In 2004, Tivey et al. designed a low-power UOWC system using an array of light-emitting diodes (LEDs), silicon-based photodetectors, and an Infrared Data Association (IrDA) protocol [24]. This proved that underwater optical wireless communication systems could transmit data using low-power, low-cost LEDs.

In 2004, Schill et al. established a small real-time optical communication transceiver using a group of small autonomous underwater vehicles. The transceiver used a 3-W high-power blue-green LED and a photodiode sensitive to wavelengths between 460 and 520 nm, achieving a data rate of 57.6 kbit/s in the 1–2-m communication range [25].

In 2008, Hanson et al. used the Monte Carlo method to simulate the transmission effect of underwater optical wireless communication systems in three different water bodies; a data transfer rate of 1 Gbit/s was realized in a 2-m water link [26]. Since the first demonstration of Gbit/s data rates, underwater optical wireless communications have explored higher data rates and longer transmission distances.

In 2013, Cossu et al., using two low-cost blue LED arrays as light emitters and an avalanche photodiode module as a receiver, made measurements using a 3.3-m-diameter outdoor water tank. It consisted of two waterproof boxes holding the water tank boundary transmitter and receiver. It achieved 2.5-m high-speed underwater wireless error-free transmissions using 6.25-Mbit/s phase coding, 12.5-Mbit/s non-return-to-zero (NRZ) 8B/10B coding, and 58-Mbit/s discrete multi-audio modulation [27].

In 2015, Nakamura et al. conducted an experiment with a 1.45-Gbit/s communication rate in a 4.8-m communication distance by using a 405-nm blue light LD and a field programmable gate array (FPGA) to generate 16-quadrature amplitude modulation (QAM)–orthogonal frequency-division multiplexing (OFDM) modulation signals in real time. Its bit error rate was 9.1×10^{-4} [28].

In 2016, Oubei et al. conducted research on a UOWC system using a 448-nm blue light LD 16-QAM-OFDM modulation scheme. A transmission rate of 3.2 Gbit/s over a transmission distance of 2.5 m was achieved by placing a reflector in a 0.6-m water tank; the bit error rate was 6.83×10^{-4} [29].

In 2019, Tsai et al. first proposed a 30-Gbit/s four-stage pulse-amplitude modulation (PAM4) underwater optical wireless communication system. The system was equipped with a beam reducer/expander on a 12.5-m pipe underwater channel and 2.5-m high-turbidity harbor underwater channel link. The large amount of scattered light received by the optical receiver of the PAM4 system improved with the use of the beam constrictor/expander [30].

In 2022, Kong et al., of King Abdullah University of Science and Technology, built a 2-K real-time digital video surveillance system using 458-nm LEDs, based on a wide-field underwater optical wireless communication system. The video stream was processed through transmitter circuits including video coding (H. 256) and on–off keying (OOK) modulation, in a 1.5-m pure water channel and 1.53-ml/s bubble-induced turbulent pure water channel. They transmitted a real-time video at a 2.5-Mbit/s communication rate [31].

From the foreign research progress of underwater optical wireless communication, it can be seen that researchers are keen to study UOWC through different modulation and coding methods, by changing the modulation mode of the UOWC system to improve the transmission rate and transmission distance, and through a variety of UOWC system analysis simulation models. By comparing simulation data, such as communication rate and communication distance, with engineering experiments, the theoretical data were verified experimentally, confirming that the models improved the communication performance of the UOWC system.

Domestic Research Status

In 2006, Zhou et al. used photoelectric detector array reception, high-speed analog-to-digital conversion, efficient data acquisition, and digital signal processing (DSP) technology to design and implement a high-sensitivity underwater receiver with a DSP processor that could receive pulse-position modulation (PPM) signals. Compared with a hard receiver, a receiver with DSP improved the sensitivity by at least 3 dB [32].

In 2008, Luo et al., using a blue-green LED as the emitting light source, designed a set of UOWC systems according to the IrDA physical protocol. The system realized underwater visible light communication and detected the light intensity using a linear distance measurement method. The transmission rate reached 125 kbit/s without the influence of any light source within 2 m. The average distance

difference was 5% when the LED light source and receiver were aligned in a straight line [33].

In 2009, Sui et al. simulated a UOWC communication channel link. By combining the Bose–Chaudhuri–Hocquenghem (BCH) error-correcting code and OOK modulation technology, the communication performance in water was analyzed through the simulation of different modulation modes. Using an LED as the light source of an underwater communication system, information was reliably transmitted when the communication transmission rate was 4 Mbit/s [34].

In 2013, Pang et al. used seven blue LEDs in series as the emitting light source and a blue light-enhanced PD as the receiving system. A set of underwater optical wireless communication systems was designed, based on the binary open–close keying principle, and a 3-Mbit/s data transmission was realized through a 13-m communication channel [35].

In 2014, Tao et al. designed and implemented an underwater optical wireless communication node and transplanted the μC/OS-II embedded operating system onto the hardware platform of the communication node. The physical layer protocol and data link layer protocol were implemented, based on the IEEE 802.15.7 visible light communication standard. The visible light communication nodes could communicate reliably within a range of 1 m under water [36].

In 2016, Xu et al. implemented an intensity modulation/direct detection (IM/DD)-OFDM-based underwater optical wireless communication system using a single compact blue LED and a low-cost PIN photodiode. The system achieved a communication rate of 161.36 Mbit/s in a 2-m underwater tank [37].

In 2017, Huang et al. limited the frequency response of an avalanche photodiode by using a GaN blue laser diode as a transmitter and 16 QAM-OFDM for data transmission, with a data rate of up to 14.8 Gbit/s. The transmission distance in seawater was 1.7 m [38].

In 2019, Hong et al. realized an underwater laser communication system with a 5-m transmission distance and a net communication rate of 18.09 Gbit/s using discrete multi-tone (DMT) modulation with probability constellation shaping (PCS) technology. This was the first time that PCS-QAM-DMT was used in an underwater laser communication system; it was also the highest data communication rate achieved using a single laser diode in a UOWC system [39].

In 2019, Wang et al. successfully implemented an underwater laser communication system with a 500-Mbit/s communication rate through a 100-m tap water channel using a 520-nm green LD and NRZ-OOK modulation of a single-mode tail fiber [40].

In 2021, using spread-spectrum technology to improve the receiver sensitivity of a UOWC system, researchers extended the transmission distance of the system. In addition, the system's anti-interference ability was improved. The receiver sensitivity of a communication system loaded with 200-MHz signals was increased by 4.8 dB [41].

In 2022, Ke et al. proposed a design scheme for a wide-angle-focused double-layer compound eye lens optical system. A compound eye lens receiving system was designed and simulated using the Zemax software, and an experimental setup for

investigating the underwater communication system receiving efficiency was built. The results showed that the double-layer compound eye lens receiving antenna significantly improved the optical power received by the receiving end [42].

Underwater optical wireless communication research can be divided into two categories: theoretical and experimental. In the theoretical aspect, the underwater channel is modeled, and the track of each photon in the water is analyzed under a certain transceiver angle and emitting light power. The photon is captured by the receiving end of the system after being repeatedly scattered by water.

The experimental aspect is focused on developing an underwater optical wireless communication system with real-time communication. Researchers mainly focus on improving the power of the transmitter light source, adopting a high-sensitivity detector and a suitable signal-modulation format, and adding a signal processing method with higher anti-noise performance to improve the communication speed and transmission distance of the underwater optical wireless communication system.

To date, researchers around the world have conducted many theoretical simulations and experimental research; however, many underwater models and communication methods built using simulations still cannot be verified experimentally—for example, considering the complexity of the water environment—owing to technical problems and the inability to improve the detector sensitivity.

Furthermore, it is necessary to experimentally obtain different water environment data to build more perfect underwater models and guide theoretical experiments. With a solution to these problems, it will be easier for underwater optical wireless communication systems to meet the needs of different water environments and contribute to existing underwater communication.

Characteristics of Seawater Channels

Seawater is a complex and irregular mixed medium in underwater transmission channels. The turbulence caused by the movement of water molecules, the chlorophyll produced by phytoplankton, yellow substances (such as fulvic and humic acid) produced by various fish and algae, and various dissolved salts and mineral particles affect the transmission characteristics of laser beams in seawater.

The random fluctuation of ocean turbulence and the absorption and scattering characteristics of underwater suspended particles attenuate the laser beam energy and cause the expansion, drift, and scintillation of the laser beam. This increases the communication difficulty and affects the system performance. Therefore, this section introduces the components of ocean channels, basic theory of ocean turbulence, attenuation characteristics of underwater suspended particles, and key technical difficulties of underwater laser transmissions.

Ocean Turbulence Spectrum Models

Since the beginning of the 1990s, different ocean turbulence spectrum models have appeared one after another. The ocean turbulence spectrum model proposed by Nikishov is the most widely used. It considers the factors of temperature and salinity and is suitable for isotropic ocean turbulence.

The expression for the ocean turbulence spectrum is [43]

$$\begin{aligned} E_n(k) = {} & C_0 \chi_n \varepsilon^{-1/3} \kappa^{-11/3} \left[1 + C_1 (\kappa\eta)^{2/3}\right] \\ & \times \frac{\omega^2 \theta \exp(-A_T \delta) + \exp(-A_S \delta) - \omega(1+\theta)\exp(-A_{TS}\delta)}{\omega^2 \theta + 1 - \omega(1+\theta)}, \end{aligned} \tag{32.1}$$

where $C_0 = 0.07$, $C_1 = 2.3$, $\chi_n = \alpha^2 \chi_T + \alpha^2/\omega^2 \chi_T - 2\alpha^2/\omega^2 \chi_T$, $\theta = K_T/K_S$, K_T denotes the turbulent heat dissipation coefficient, and K_S denotes the turbulent salinity diffusion coefficient. χ_n is substituted into

$$\begin{aligned} E_n(\kappa) = {} & C_0 \left(\alpha^2 \chi_T + \frac{\alpha^2}{\omega^2}\chi_T - \frac{2\alpha^2}{\omega}\right) \varepsilon^{-1/3} \kappa_{-11/3} \left[1 + C_1 (\kappa\eta)^{2/3}\right] \\ & \times \frac{\omega^2 \theta \exp(-A_T \delta) + \exp(-A_S \delta) - \omega(1+\theta)\exp(-A_{TS}\delta)}{\omega^2 \theta + 1 - \omega(1+\theta)}. \end{aligned} \tag{32.2}$$

For isotropic ocean turbulence, $K_T = K_S$, $\theta = 1$, and the scalar spectral model of ocean turbulence in Eq. (32.2) can be simplified to [43]

$$\begin{aligned} E_n(k) = {} & C_0 \alpha^2 \varepsilon^{-1/3} \kappa^{-11/3} \left[1 + C_1 (\kappa\eta)^{2/3}\right] \chi_T/\omega^2 \\ & \times \omega^2 \exp(-A_T \delta) + \exp(-A_S \delta) - 2\omega \exp(-A_{TS}\delta). \end{aligned} \tag{32.3}$$

According to the relationship between the scalar spectrum and the power spectrum function, the power spectrum function $\phi_n(\kappa)$ in ocean turbulence can be written as

$$\begin{aligned} \phi_n(\kappa) = {} & (4\pi)^{-1} \kappa^{-2} E_n(\kappa) \\ = {} & 0.388 \times 10^{-8} \varepsilon^{-1/3} \kappa^{-11/3} \left[1 + 2.35 (\kappa\eta)^{2/3}\right] \frac{\chi_T}{\omega^2} \\ & \times \left[\omega^2 \exp(-A_T \delta) + \exp(-A_S \delta) - 2\omega \exp(-A_{TS}\delta)\right], \end{aligned} \tag{32.4}$$

where ε is the turbulent kinetic energy dissipation rate with a value range of $(10^{-1} - 10^{-10})\mathrm{m}^2/\mathrm{s}^3$ and the Kolmogorov-scale efficiency η has a value of $(6 \times 10^{-5} - 0.01)\mathrm{m}$. The temperature variance dissipation rate χ_T has a value range of $(10^{-4} - 10^{-10})\mathrm{K}^2/\mathrm{s}$ from the sea surface to the deep sea. In the temperature induction to salinity induction ratio $\omega = [-5, 0]$, $\omega = 0$ indicates that salinity changes play a major role; $\omega = -5$ indicates that temperature changes play a major role. Other parameters

include $A_T = 1.863 \times 10^{-2}$, $A_S = 1.9 \times 10^{-4}$, $A_{TS} = 9.41 \times 10^{-3}$, and $\delta = 8.284(\kappa\eta)^{4/3} + 12.978(\kappa\eta)^2$.

The expression for the power spectrum $\phi_n(\kappa_x, \kappa_y)$ of the refractive index of anisotropic ocean turbulence is

$$\begin{aligned}\phi_n(\kappa_x, \kappa_y) = {} & 0.388 \times 10^{-8} \mu_x \mu_y \varepsilon^{-1/3} \left[(\mu_x \kappa_x)^2 + (\mu_y \kappa_y)^2\right]^{-11/6} \\ & \left\{1 + 2.35\left[(\mu_x \kappa_x)^2 + (\mu_y \kappa_y)^2\right]^{1/3} \eta^{2/3}\right\}(\chi_T/\omega^2) \\ & \times \left[\omega^2 \exp(-A_T\delta) + \exp(-A_S\delta) - 2\omega \exp(-A_{TS}\delta)\right],\end{aligned} \tag{32.5}$$

where μ_x and μ_y are, respectively, the anisotropy factors of the x and y directions; the meanings of parameters ε, η, χ_T and ω are shown above; the expression is

$$\delta = 8.284\eta^{4/3}\left[(\mu_x \kappa_x)^2 + (\mu_y \kappa_y)^2\right]^{2/3} + 12.978\eta^2\left[(\mu_x \kappa_x)^2 + (\mu_y \kappa_y)^2\right]. \tag{32.6}$$

Attenuation Characteristics of Underwater Particles

Absorption Properties

When light travels through seawater, it is absorbed by water molecules, plankton, and algae; this is known as the absorption characteristic of seawater. The factors that affect the absorption properties of seawater can be grouped into four categories: seawater itself, chlorophyll-containing particles (phytoplankton), suspended particles in seawater, and colored dissolved organic matter. The combination of these four absorption types constitutes the absorption of light by seawater. Therefore, the absorption and attenuation coefficient of seawater can be expressed as [44]

$$\alpha(\lambda) = \alpha_w(\lambda) + \alpha_{\text{chl}}(\lambda) + \alpha_{\text{nap}}(\lambda) + \alpha_y(\lambda), \tag{32.7}$$

where $\alpha_w(\lambda)$ denotes the absorption and attenuation coefficient of pure water, $\alpha_{\text{chl}}(\lambda)$ denotes that of chlorophyll-containing particles (phytoplankton), $\alpha_{\text{nap}}(\lambda)$ denotes that of suspended particles, and $\alpha_y(\lambda)$ represents that by yellow matter. In the same area, when the salinity, depth, and measurement time change, the absorption characteristics of the seawater also change. These four effects are described below.

Light Absorption by Pure Water

The absorption of light by seawater is mainly caused by water molecules and dissolved salts. The concentration of inorganic salts in seawater is approximately 3.8%, including NaCl, KCl, and $CaCl_2$. A large number of experiments have shown that the effect of dissolved salts in seawater can be disregarded because of their low light absorption compared to that of water molecules [45]. Therefore, the absorption of light by seawater is primarily caused by the water molecules.

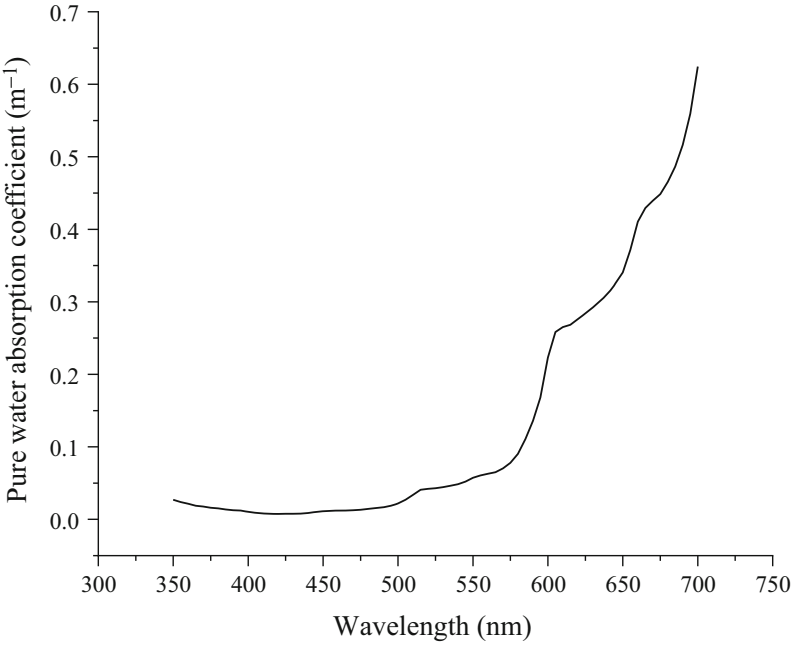


Fig. 32.1 Absorption coefficient of pure seawater at different wavelengths [47]

Morel et al. measured a large amount of pure seawater using a spectrophotometer. They first measured the attenuation coefficient of a pure water beam and then subtracted the scattering coefficient from the data to calculate the absorption coefficient of pure water to light. The relationship between the absorption coefficient $\alpha_w(\lambda)$ of pure seawater at different wavelengths and the light wavelength λ is obtained [46], as shown in Fig. 32.1.

As shown in Fig. 32.1, the absorption coefficient of pure water is minimal when the wavelength of light is between 350 and 550 nm, which is in the blue-green band. With an increase in the wavelength, the absorption coefficient of pure water increases gradually, with small shoulder peaks at 510 and 600 nm, and then increases rapidly.

Light Absorption Characteristics of Phytoplankton

Similar to plants on land, phytoplankton in seawater use chlorophyll for photosynthesis; therefore, their ability to absorb light waves depends on the amount of chlorophyll they contain. According to Prieur et al., different types of chlorophyll have different light absorption capacities at different wavelengths, and the ratio of chlorophyll a in phytoplankton is the highest. Therefore, the absorption coefficient of phytoplankton to light can be determined by the concentration of chlorophyll a. The absorption coefficient of chlorophyll a to light waves is [47]

$$\alpha_{\mathrm{chl}}(\lambda) = a_c{}^0(\lambda)\left(\frac{C_c}{C_c{}^0}\right)^{0.602}, \tag{32.8}$$

where C_c is the concentration of chlorophyll, $C_c{}^0 = 1$ mg/m^3, and $a_c{}^0(\lambda)$ is the absorption coefficient of the chlorophyll ratio in phytoplankton.

In this study, the absorption coefficient of the chlorophyll ratio at different wavelengths of light $a_c{}^0(\lambda)$ in reference [48] was substituted into Eq. (32.8). The absorption coefficient of chlorophyll at different light wavelengths $\alpha_{\mathrm{chl}}(\lambda)$ was calculated when the chlorophyll concentration was 1 mg/m^3. The relationship between the absorption coefficient of phytoplankton $\alpha_{\mathrm{chl}}(\lambda)$ and the wavelength of light λ is shown in Fig. 32.2.

As shown in Fig. 32.2, when the wavelength of light is between 550 and 650 nm, the absorption coefficient of chlorophyll a of phytoplankton is the smallest, with two peaks at wavelengths of 440 and 670 nm. This suggests that the chlorophyll of phytoplankton has the greatest absorption coefficient for blue-violet and orange-red light.

Light Absorption Properties of Yellow Matter

Yellow matter in seawater refers to the colored dissolved organic matter formed by humus from animal and plant decay; it is called yellow matter because of its yellow-brown color. Experimental measurements by Prieur et al. in 1977 showed that the

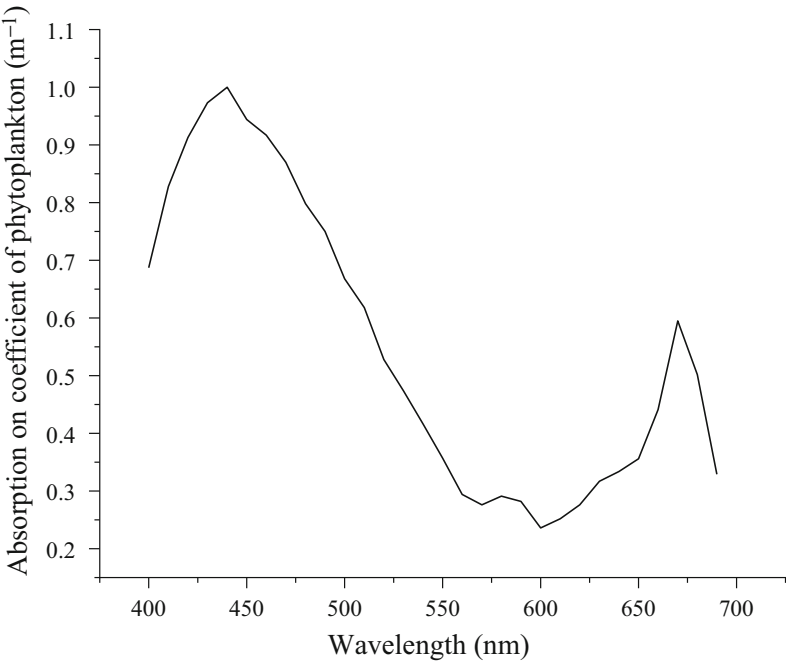


Fig. 32.2 Absorption coefficient of phytoplankton at different wavelengths [47]

decay of light absorbed by yellow matter in seawater is exponential [48]. In 2002, Wu et al. derived a mathematical model of the absorption characteristics of yellow matter during the transmission of light through seawater using experimental measurements [49]:

$$\alpha_y(\lambda) = \alpha_y(\lambda_0) \times \exp\left[-S_y \times (\lambda - \lambda_0)\right], \tag{32.9}$$

where $\alpha_y(\lambda_0)$ is the reference wavelength λ_0, the absorption rate of yellow matter in seawater is a constant, and S_y is the exponential slope of the absorption spectrum curve of yellow matter.

In 1977, Morel et al. provided an exponential slope parameter S_y value of 0.0140 nm^{-1} [48]. In 1981, Bricaud et al. collected 105 water samples from different waters and obtained a range of S_y values of 0.0100–0.0200 nm^{-1} [50]. Pegau et al. (1995) obtained a mean S_y value of 0.0170 nm^{-1} [51] by measuring 26 lake water samples. Wu et al. (2002) collected samples from waters, such as Jiaozhou Bay, and obtained a range of S_y values of 0.0131–0.0180 nm^{-1} [49].

The absorption coefficient of yellow matter to light is calculated with reference to the 2002 measurement data of Wu et al. Here, $\lambda_0 = 400$, $\alpha_y(400) = 0.650\ m^{-1}$, and the exponential slope $S_y = 0.0141$. The absorption coefficient of yellow matter to light is calculated, and the relationship between the absorption coefficient of yellow matter $\alpha_y(\lambda)$ and the wavelength of light λ is plotted, as shown in Fig. 32.3.

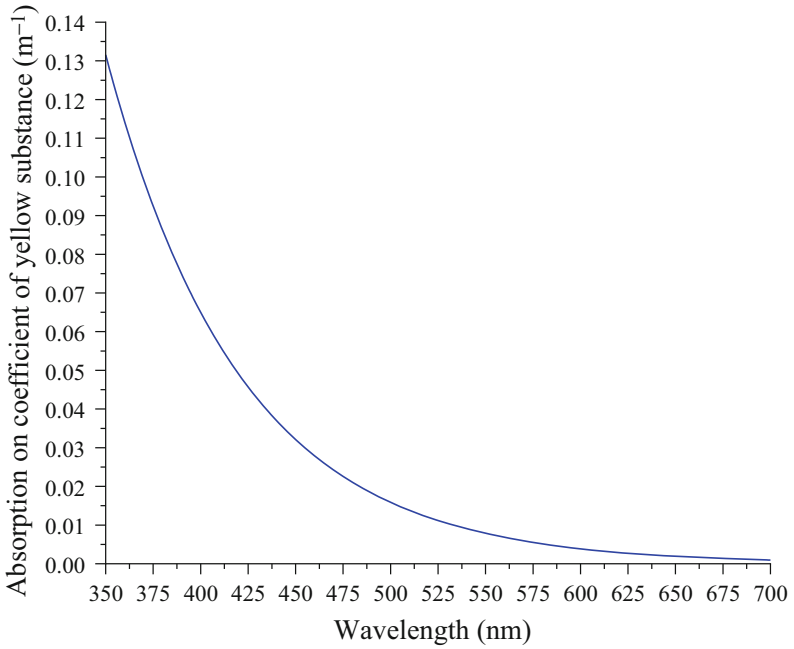


Fig. 32.3 Relationship between yellow matter and the wavelength [49]

As shown in Fig. 32.3, in the visible light band between 350 and 700 nm, the absorption spectrum of yellow matter in seawater shows an exponential change. With an increase in wavelength, a monotonically decreasing trend is observed.

Light Absorption Characteristics of Suspended Particles

Suspended particles in seawater are matter particles between a few and a few hundred microns in size, mainly biological waste, debris, and suspended sediment particles, which also absorb light waves. Owing to the insolubility of suspended particles, they are mainly distributed in coastal areas affected by river discharge and coastal erosion. The total amount of suspended particles is small in deep-sea areas far from the coast, owing to the low particle deposition and influence of human factors.

Gilerson et al. conducted field measurements at 42 sites along the Maryland coast and adjacent rivers and ports in the Chesapeake Bay. The absorption spectra of suspended particles in seawater measured by the ocean optical spectrometer were the same as those of yellow matter, and both were modeled as exponential decreases with the increase of the wavelength.

The functional expression for the absorption characteristics of suspended particles in seawater is [52]

$$\alpha_{\text{nap}}(\lambda) = \alpha_{\text{nap}}(\lambda_0) \times \exp\left(-S_{\text{nap}}(\lambda - \lambda_0)\right) \tag{32.10}$$

$$\alpha_{\text{nap}}(\lambda_0) = C_{\text{nap}} \times \alpha^*_{\text{nap}}(\lambda_0), \tag{32.11}$$

where $\alpha^*_{\text{nap}}(\lambda_0)$ is the specific absorptivity of suspended particles at the reference wavelength λ_0, C_{nap} is the concentration of non-algal suspended particles, and S_{nap} is the exponential slope of the spectral curve of light absorption of suspended particles.

According to the data of Gilerson et al. in 2007 [52], we calculate the absorption coefficient of light using Eq. (32.10), where the reference wavelength $\lambda_0 = 400$; then, $\alpha^*_{\text{nap}}(400) = 0.0375\ \text{m}^2/\text{g}$, $C_{\text{nap}} = 10\ \text{g/m}^3$, and the exponential slope $S_{\text{nap}} = 0.015$. The absorption coefficient $\alpha_{\text{nap}}(\lambda_0)$ of the suspended particles is calculated by simulation, and the relationship between the absorption coefficient $\alpha_{\text{nap}}(\lambda)$ and the wavelength λ is shown in Fig. 32.4.

As shown in Fig. 32.4, in the visible light band between 350 and 700 nm, the light absorption spectrum of suspended particles in seawater presents an exponential change form and with the increase in wavelength presents a monotonically decreasing trend. When the light wavelength is greater than 650 nm, the light absorption of the suspended particles approaches zero [52].

Scattering Characteristics

In addition to the loss caused by the absorption characteristics of seawater, a large number of particles and organic matter are found near land water, where the attenuation of seawater is dominated by scattering; at this point, the water's minimum attenuation window moves from the blue band (470 nm) to the green band (550 nm). The presence of water molecules in seawater, chlorophyll-containing

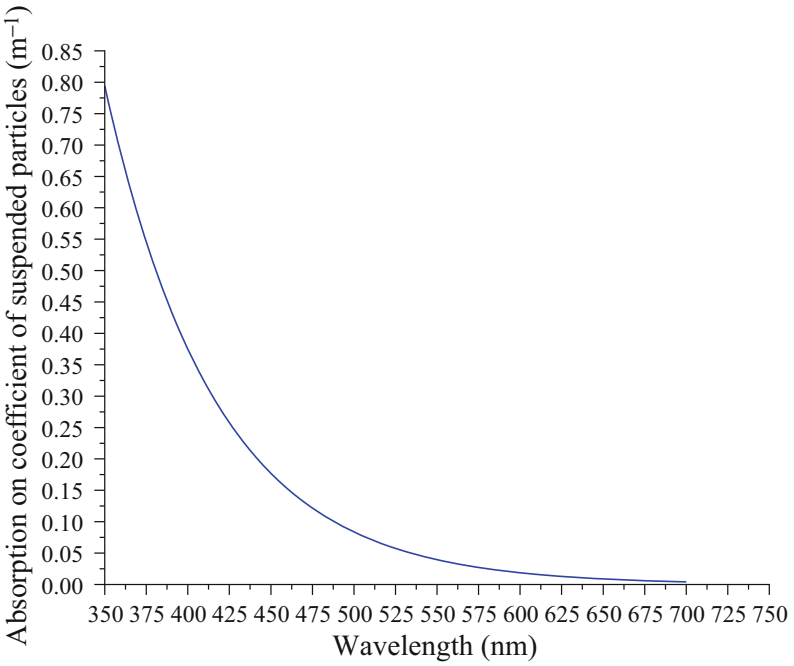


Fig. 32.4 Relationship graph of the absorption coefficient of suspended particles with different wavelengths

particles (phytoplankton), and suspended particles of different sizes also cause the light transmission to lose energy.

The combination of these three types of matter makes up the scattering of light by seawater. Therefore, the attenuation coefficient $\beta(\lambda)$ of light scattering by these three types of material can be expressed as [53]

$$\beta(\lambda) = \beta_w(\lambda) + \beta_{\text{chl}}(\lambda) + \beta_{\text{nap}}(\lambda), \tag{32.12}$$

where $\beta_w(\lambda)$ denotes the attenuation coefficient of light scattering by pure seawater, $\beta_{\text{chl}}(\lambda)$ denotes that by plankton, and $\beta_{\text{nap}}(\lambda)$ denotes that by suspended particles. These three types of effects are explained separately below.

Light-Scattering Properties of Pure Water

Water molecules in seawater absorb light and have scattering characteristics that follow the Rayleigh distribution. In 1977, Morel experimentally determined the equation of a light-scattering model using pure water [54]. The Rayleigh scattering coefficient $\beta_w(\lambda)$ of pure seawater is

$$\beta_w(\lambda) = 0.005826 \times \left(\frac{400}{\lambda}\right)^{4.322}. \tag{32.13}$$

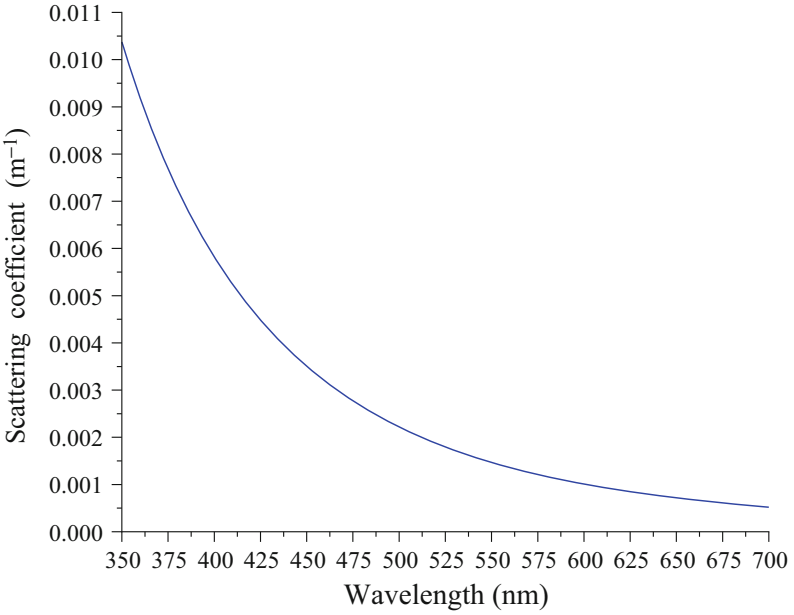


Fig. 32.5 Relationship between the scattering coefficients of pure seawater and different wavelengths

By substituting the wavelength λ (350–700 nm) of the visible light band into Eq. (32.13), the relationship between the scattering coefficient $\beta_w(\lambda)$ of pure seawater and that of the visible light band λ (350–700 nm) can be obtained. As shown in Fig. 32.5, the scattering coefficient $\beta_w(\lambda)$ of pure seawater decreases exponentially with the increase of light wavelength.

Phytoplankton Light-Scattering Characteristics

The light-scattering characteristics of phytoplankton are also related to their chlorophyll concentrations. To simulate the scattering of phytoplankton, Roesler et al. obtained the equation for scattering coefficient $\beta_{\text{chl}}(\lambda)$ of phytoplankton using the Mie theory and field measurements [54]:

$$\beta_{\text{chl}}(\lambda) = 0.45\left(\frac{550}{\lambda}\right)C^{0.62}, \tag{32.14}$$

where C is the concentration of phytoplankton chlorophyll.

Carina et al. optically simulated the change in the carbon concentration in the core by changing the real part of the refractive index of the particle and obtained a chlorophyll backscattering ratio of [55]

$$\beta_{\text{pc}} = 2\pi\chi_p\left(124^\circ\right)\beta_p\left(124^\circ\right), \tag{32.15}$$

where $\chi_p(124^\circ)$ and $\beta_p(124°)$ are the ratio constants, $\chi_p(124^\circ) = 1.076$, and $\beta_p(124°) = 0.00076$. Thus, the backscattering coefficient of chlorophyll to light is

$$\beta_{\mathrm{cb}}(\lambda) = \beta_{\mathrm{chl}}(\lambda) \times \beta_{\mathrm{pc}}. \tag{32.16}$$

The forward-scattering coefficient of chlorophyll is

$$\beta_{\mathrm{cf}}(\lambda) = \beta_{\mathrm{chl}}(\lambda) - \beta_{\mathrm{cb}}(\lambda). \tag{32.17}$$

By taking the chlorophyll concentration values of different phytoplankton, $C = 0.01$ mg/m^3, 0.05 mg/m^3, 0.1 mg/m^3, 0.5 mg/m^3, and 1 mg/m^3, the backscattering coefficient $\beta_{\mathrm{cf}}(\lambda)$ and the forward-scattering coefficient $\beta_{\mathrm{cb}}(\lambda)$ of phytoplankton to light were obtained using Eqs. (32.16) and (32.17), as shown in Fig. 32.6.

As can be seen from Fig. 32.6, the light-scattering coefficient $\beta_{\mathrm{chl}}(\lambda)$ of phytoplankton decreases with the increase of wavelength. While the light wavelength remains constant, the scattering coefficient of phytoplankton increases with the increase of the chlorophyll concentration.

Suspended Particle Light-Scattering Characteristics

Particles suspended in seawater are mainly composed of biological debris, excreta, decomposed matter, suspended sediment particles, etc. However, because the shape and size of suspended particles in seawater are difficult to determine, the scattering model of suspended particles is simplified by the Mie theory and the Rayleigh scattering model to facilitate the calculation of the scattering coefficient.

The suspended particles are divided into small and large particles according to their diameter. The scattering coefficient equation of the suspended particles is [56]

$$\beta_{\mathrm{nap}}(\lambda) = \beta^0_{\mathrm{snap}}(\lambda) C_s + \beta^0_{\mathrm{lnap}}(\lambda) C_l, \tag{32.18}$$

where $\beta^0_{\mathrm{snap}}(\lambda)$ is the specific scattering coefficient of small particles, $\beta^0_{\mathrm{lnap}}(\lambda)$ is the specific scattering coefficient of large particles, C_s is the concentration of small particles, and C_l is the concentration of large particles.

Specific Light-Scattering Properties of Small Particles

Particles suspended in seawater that are smaller than 5 μm in diameter are called small particles; they have a density of 2 g/cm^3 and a specific light-scattering coefficient $\beta^0_{\mathrm{snap}}(\lambda)$. It can be derived from the spectral correlation of small suspended particles as follows [53]:

$$\beta^0_{\mathrm{snap}}(\lambda) = 1.151302 \left(\frac{400}{\lambda}\right)^{1.7}. \tag{32.19}$$

In this equation, the concentration C_s of small suspended particles is mainly affected by the concentration C_c of chlorophyll in seawater. The expression of the correlation is [53]

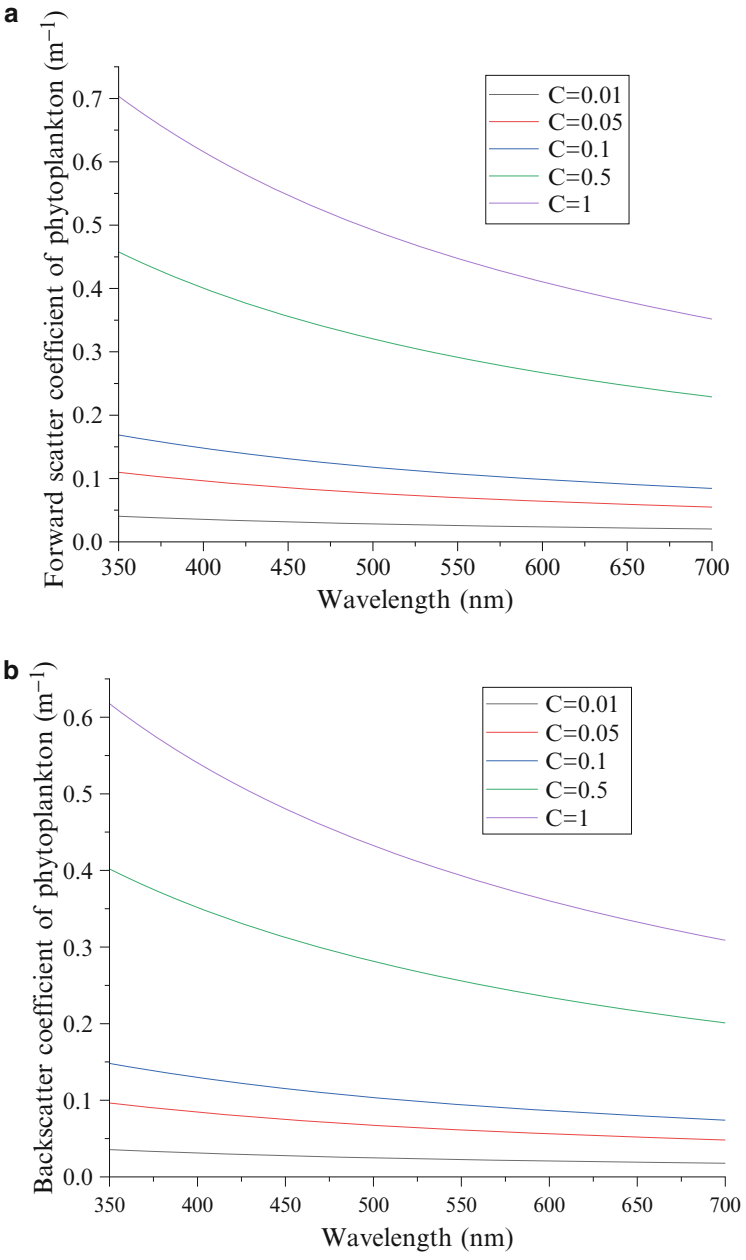


Fig. 32.6 Relationship between the scattering coefficient of phytoplankton and the light wavelength. (**a**) Forward scattering; (**b**) backscattering

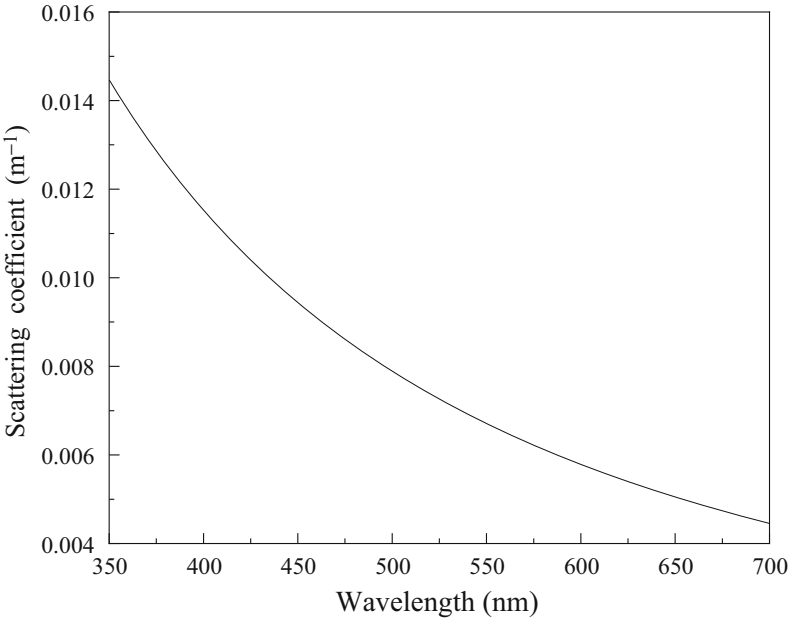


Fig. 32.7 Scattering coefficient of small suspended particles with different wavelengths

$$C_S = 0.01739C \exp\left[0.11631\left(\frac{C}{C_c^0}\right)\right], \tag{32.20}$$

where C is the chlorophyll concentration of suspended particles and ${C_c}^0$ is the chlorophyll concentration in seawater. Because Eq. (32.20) can calculate the concentration C_s of small particles when the chlorophyll concentration is ${C_c}^0 = 1$ mg/m^3, we substitute Eq. (32.20) for the visible light band λ (350–700 nm) to obtain the light-scattering coefficient of small particles, as shown in Fig. 32.7.

As shown in Fig. 32.7, the scattering properties of suspended particles decrease with an increase in wavelength; however, the scattering coefficient of small suspended particles changes very little over the entire visible wavelength range, decreasing only by 0.009 m^{-1}.

Light-Scattering Characteristics of Large Particles

Suspended particles larger than 5 μm in diameter are called large particles. The density of large particles is 1 g/cm^3, and the refractive index is 1.03. The specific scattering coefficient of large particles is $\beta^0_{\text{lnap}}(\lambda)$; it can be derived from the spectral correlation of large suspended particles as [53]

$$\beta^0_{\text{lnap}}(\lambda) = 0.341074\left(\frac{400}{\lambda}\right)^{0.3}. \tag{32.21}$$

In the equation, the concentration C_l of large suspended particles is mainly affected by the concentration C_c of chlorophyll in seawater. The expression of the correlation is [53]

$$C_l = 0.76284C \exp\left[0.03092\left(\frac{C}{C_c^0}\right)\right], \tag{32.22}$$

where C is the chlorophyll concentration of the suspended particles and ${C_c}^0$ is the chlorophyll concentration of seawater. The C_l concentration of large particles can be calculated as $C_c^0 = 1\,\mathrm{mg/m^3}$ from the Eq. (32.22). Hence, we substitute Eq. (32.22) for λ (350–700 nm) in the visible band to obtain the light-scattering coefficient of the large particles, as shown in Fig. 32.8.

As shown in Fig. 32.8, the scattering properties of large suspended particles in seawater decrease with an increase in the light wavelength. Compared with the scattering coefficients of small particles, when the wavelength of light is fixed, the scattering properties of large suspended particles in seawater decrease with an increase in the light wavelength, as shown in Fig. 32.8. When the light wavelength is constant, the scattering effect of large suspended particles in seawater is significantly higher than that of small suspended particles.

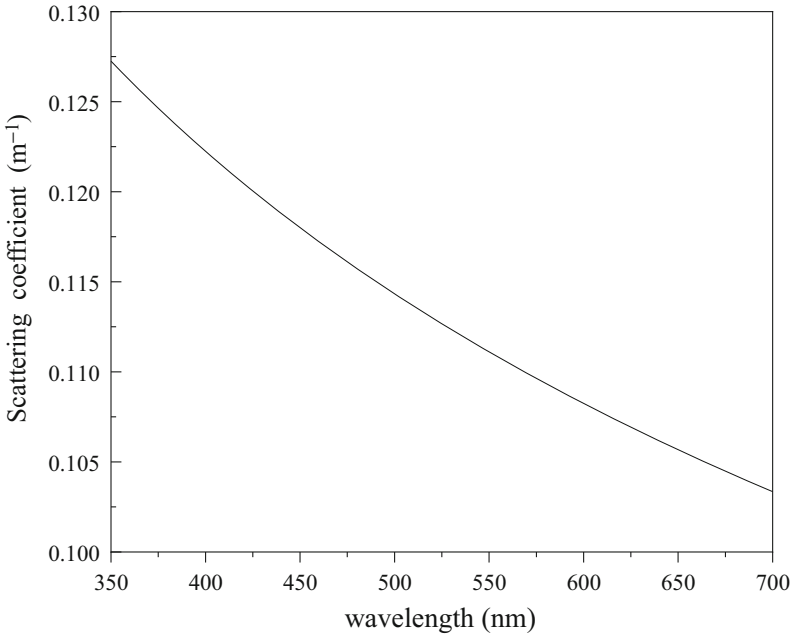


Fig. 32.8 Plot of the scattering coefficient of large suspended particles with different wavelengths

Experimental Measurements of Underwater Optical Wireless Communication

The effect of underwater turbulence on laser transmissions leads to a significant reduction in the accuracy and transmission distance of laser communication and may even interrupt communications. In this section, the Chan River, Ba River, and Yantai No. 1 beach in Xi'an, Shaanxi Province, were surveyed. The skewness and steepness, underwater structure constant, probability density distribution, scintillation index, optical power loss, and bit error rate (BER) of the underwater optical signal samples at different water velocities and distances were analyzed.

Experimental System

The experimental system is shown in Fig. 32.9. The emitting end had a 532-nm wavelength PL-E Pro green laser; its output power was 1 W, and the operating voltage was 4.2 V. The detector was a PD300-UV photodiode sensor with a spectral response range of 200–1100 nm and aperture size of 10 mm × 10 mm. The filter had a wavelength range of 513–545 nm and aperture of 20 mm. In addition, the underwater temperature was measured using a probe-type electronic thermometer and the current water velocity was averaged using the floating method.

After the laser was transmitted through an underwater channel, the emission of the green laser was analyzed, first, through a double-layer compound eye lens [57]. A curved lens on the second layer of the double-layer compound eye lens converged and superposed the light beams converged by the lens array, so that the light passing through each small lens was merged onto the detector surface. The detector and filter were then used to measure the instantaneous optical signal, which was converted into an electrical signal and sent to a computer for data acquisition and signal processing.

The experimental system is shown in Fig. 32.9. In this experiment, the transmitter and receiver were first aligned. The laser beam was passed through a double-layer compound eye lens. When the laser was stable, the PD300-UV detector was used to collect the optical signals. The temperature was recorded simultaneously. Finally, the light power of the light signal was analyzed, and the fluctuation characteristics of the light signal sample and light power loss under different turbulence intensities were obtained.

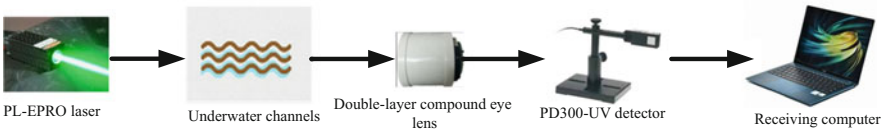


Fig. 32.9 Experimental system diagram

Table 32.1 Basic experimental conditions at different water velocities

Location	Measurement sample	Date	Link distance (m)	Wind speed (m/s)	Water temperature (° C)	Water speed (m/s)
Chan River	T1	2023.2.17	0.5	2	5–5.2	0.34
	T2	2023.2.17	0.5	2	5.2–5.4	0.47
	T3	2023.2.17	0.5	2	5.4–5.6	0.53
	T4	2023.2.17	0.5	2	5.6–5.8	0.64
Ba River	T5	2023.3.5	1.5	2	7.9–8.1	0.43
	T6	2023.3.5	1.5	2	8.1–8.3	0.52
	T7	2023.3.5	1.5	2	8.3–8.5	0.61
	T8	2023.3.5	1.5′	2	8.5–8.7	0.70

Table 32.2 Different basic conditions of the water downlink experiment

Location	Measurement sample	Date	Link distance (m)	Wind speed (m/s)	Water temperature (° C)	Water direction
Chan River	T9	2023.2.19	0.5	2	5–5.2	Smooth sailing
	T10	2023.2.19	0.5	2	5.2–5.4	Adverse water
	T11	2023.2.19	0.5	2	5.4–5.6	Horizontal
Ba River	T12	2023.3.6	1.5	1	8–8.2	Smooth sailing
	T13	2023.3.6	1.5	1	8.2–8.4	Adverse water
	T14	2023.3.6	1.5	1	8.4–8.6	Horizontal

Table 32.3 Basic conditions of experiments at different distances

Location	Measurement sample	Date	Link distance (m)	Wind speed (m/s)	Water temperature (° C)
Yantai No. 1 Beach	T15	2023.4.2	1	3.3	8.5–8.8
	T16	2023.4.2	3	3.3	8.8–9.1
	T17	2023.4.2	5	3.3	9.1–9.4
	T18	2023.4.2	6.5	3.3	9.4–9.7

From November 2022 to June 2023, the research group conducted underwater laser transmission experiments in the Chan River, Ba River, and Yantai No. 1 beach in Shaanxi Province. The link consisted of a transmitter, channel, and receiver. During the experiment, the data parameters were measured, as shown in Tables 32.1, 32.2, and 32.3. To eliminate the error caused by wind speed on the measured results, the experimental links were maintained 10 cm from the water surface [58]. In the measured temperature range, the variance of the sample was almost constant; therefore, the effect of temperature on the results can be ignored in the following experiments.

Statistical Analysis of Data

The moment estimation estimates the total moment by the sample moment. The first-order origin moment of the measured signal sequence X represents the expectation of the measured signal sequence, and the second-order central moment represents the variance. The skew direction and degree of skew distribution of the skew statistics data are expressed by the third-order center distance and the 3/2 power of the sample variance. The expression is [59]

$$S_X = \left\langle (\mathrm{X} - \langle X \rangle)^3 \right\rangle / \sigma_X^3. \tag{32.23}$$

The steepness measures the amount of data clustered in the center, expressed by the fourth-order center moment and variance, with the following expression [59]:

$$K_X = \left\langle (\mathrm{X} - \langle X \rangle)^4 \right\rangle / \sigma_X^4. \tag{32.24}$$

Figure 32.10 shows the relationship between the skewness and steepness of the measured samples in different waters and the variance of the measured light signal samples in horizontal coordinates. The longitudinal coordinates are the skewness and steepness of the measured light signal samples in different waters.

Figure 32.10a, b show the statistical distribution of light signal samples at different water velocities in the Chan and Ba Rivers, respectively, and Fig. 32.10c shows the statistical distribution of light signal samples at different distances from the Yantai No. 1 beach. In Fig. 32.10a–d, the sample variance of the measured optical signal increases with increasing water speed, and in graphs Fig. 32.10e, f, the sample variance of the measured optical signal increases with increasing distance. In different waters, the range of the variance distribution of the optical signal samples differs, which leads to an increase in skewness and steepness.

In Fig. 32.11, the maximum values of the skewness and kurtosis of the Chan River optical signal sample are 1.362 and 3.5, respectively, the maximum values of the Ba River optical signal sample are 1.43 and 3.8, respectively, and the maximum values of the light signal samples from the Yantai No. 1 beach are 1.49 and 4.3, respectively. The Chan River light sample had the smallest skewness and kurtosis, the Ba River light sample had the second largest skewness and kurtosis, and the Yantai No. 1 beach had the largest skewness and kurtosis.

The skewness of the optical signal sample increases with an increase in its variance, which means that the symmetry of the probability distribution of the power value of the optical signal sample worsens with an increase in the water speed and distance, compared with the normal distribution. The steepness of the optical signal samples increases gradually; however, their power distribution is increasingly dispersed, compared to that of the normal distribution.

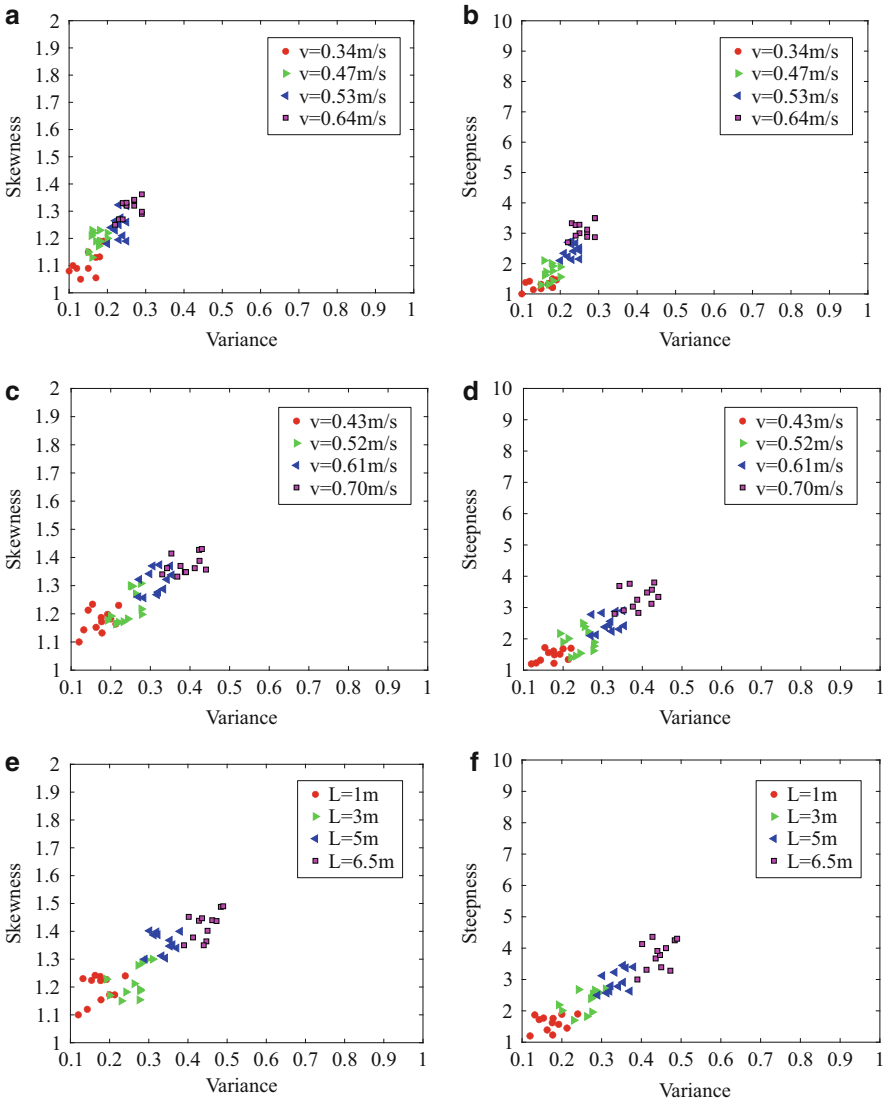


Fig. 32.10 Relationship between the deviation, steepness, and variance of measured samples in different waters. (**a**) Chan River skewness. (**b**) Chan River steepness. (**c**) Ba River skewness. (**d**) Ba River steepness. (**e**) Yantai No. 1 beach skewness. (**f**) Yantai No. 1 beach steepness

Underwater Structure Constant

The underwater refractive index structure constant C_n^2 quantitatively describes the intensity of optical turbulence. Baykal expresses the atmospheric refractive index structure constant using the ocean turbulence parameters through the modified Rytov theory. The expression is [60]

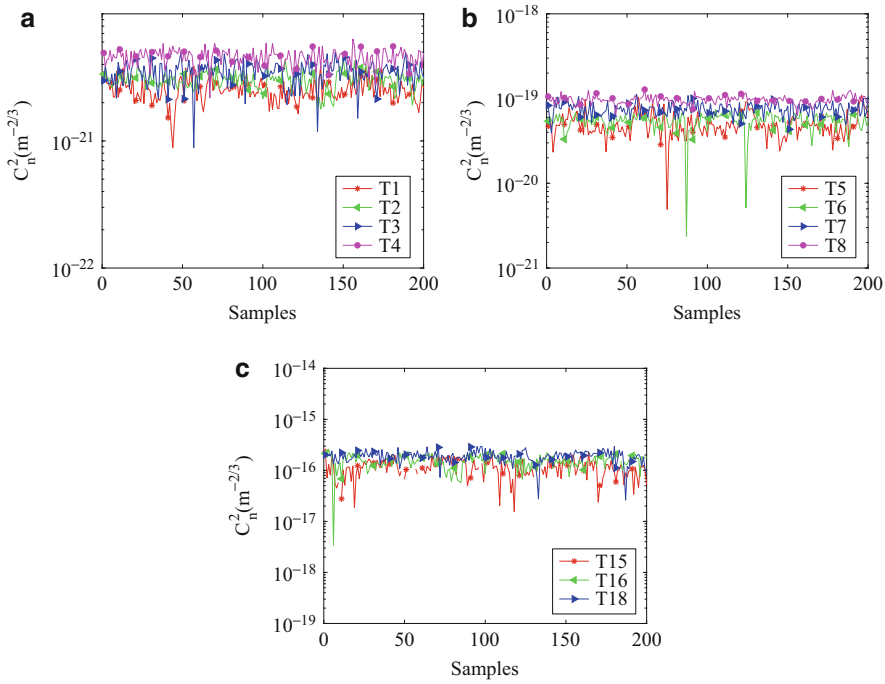


Fig. 32.11 Structure constants of light signal samples measured in different waters. (**a**) Chan River. (**b**) Ba River. (**c**) Yantai No. 1 beach

$$C_n^2 = 16\pi^2 k^{-7/6} L^{-11/6} \int_0^L \int_0^\infty \kappa \phi_n(\kappa) \cdot \left\{1 - \cos\left[\kappa^2 \zeta (L-\zeta)/(Lk)\right]\right\} d\kappa d\zeta, \tag{32.25}$$

where ζ is the distance between the Gaussian beam and the lens at the emitting end of the laser, $k = 2\pi/\lambda$ is the wave number, and λ is the wavelength. The spectral density function $\phi_n(\kappa)$ of the refractive index is expressed as

$$\Phi_n(\kappa) = 0.388 \times 10^{-8} \varepsilon^{-\frac{1}{3}} \kappa^{-\frac{11}{3}} \left[1 + 2.35(\kappa\eta)^{\frac{2}{3}}\right] \times \frac{\chi_T}{\omega^2} \times \left(\omega^2 e^{-A_T \delta} + e^{-A_S \delta} - 2\omega e^{-A_{TS} \delta}\right), \tag{32.26}$$

where the kinetic energy dissipation rate ε is defined as the rate at which the kinetic energy per unit mass of turbulence is converted into molecular thermal kinetic energy, with a typical value of 10^{-10}–10^{-1} m^2/s^3. $\delta = 1.5{C_1}^2(\kappa\eta)^{\frac{4}{3}} + {C_1}^3(\kappa\eta)^2$. The mean-square temperature dissipation rate ω is defined as the amount of turbulence acting on the temperature field through molecular heat conduction. The typical

value is 10^{-10}–10^{-4} K^2/s. The temperature–salinity contribution ratio is typically defined as the ratio of temperature-induced ocean turbulence to salinity-induced ocean turbulence, typically [−5, 0]. κ represents the magnitude of the spatial frequency, $A_T = 1.863 \times 10^{-2}$, $A_s = 1.9 \times 10^{-4}$ and $A_{TS} = 9.41 \times 10^{-3}$.

According to Davis' theory, C_n^2 can be divided into three types: strong turbulence in $C_n^2 \geq 2.5 \times 10^{-13} m^{-2/3}$, moderate turbulence in $6.4 \times 10^{-17} m^{-2/3} \leq C_n^2 \leq 2.5 \times 10^{-13} m^{-2/3}$, and weak turbulence in $C_n^2 \leq 6.4 \times 10^{-17} m^{-2/3}$. Its value changes with the water body, water speed, etc.; this section uses the scintillation method to measure the constant [61], as follows:

$$C_n^2 = 4.48 C D_t^{7/3} L^{-3} \sigma_x^2, \tag{32.27}$$

where L is the link length; $C = D_r/D_t$, where D_r and D_t are the sizes of the receiving and sending apertures, respectively; the logarithmic amplitude variance is σ_x^2, which can be expressed as [62]

$$\sigma_x^2 = \frac{1}{4} \ln\left[1 + \left(\frac{s}{< I >}\right)^2\right], \tag{32.28}$$

where the formula is the mean $< I >$ and standard deviation s of the light intensity.

To obtain the change rule C_n^2 for different water velocities and distances, the received optical signal was normalized and divided into 200 groups. By calculating the mean and standard deviation of each group of 500 light intensity values, the change curve C_n^2 was obtained using Eq. (32.27) and then Eq. (32.25), as shown in Fig. 32.11.

The maximum C_n^2, minimum C_n^2, and mean C_n^2 are calculated simultaneously to quantitatively describe the rule of taking C_n^2 values at different water velocities and distances, as shown in Table 32.4.

In experiments T1–T4, the Chan River measured samples, and C_n^2 increased with the increase of water velocity; in experiments T5–T8, the Ba River measured samples, and C_n^2 increased with the increase of water velocity; in experiments T15, T16, and T18, the Yantai No. 1 beach measured samples, and C_n^2 increased with the increase of the distance. That is, the water speed increased as the link distance increased.

Considering Fig. 32.11 and Table 32.4, it can be seen that the C_n^2 of the Chan River spanned two orders of magnitude and fluctuated randomly for different water velocity variations; most of them were distributed between 1.16×10^{-21} $m^{-2/3}$ and 5.75×10^{-21} $m^{-2/3}$. The Ba River C_n^2 was divided into three orders of magnitude, that is, between 2.3715×10^{-21} $m^{-2/3}$ and 1.5142×10^{-19} $m^{-2/3}$, and fluctuated randomly; most values were between 5.65×10^{-20} $m^{-2/3}$ and 1.21×10^{-19} $m^{-2/3}$. The Yantai No. 1 beach had four orders of magnitude, between 5.5792×10^{-19} $m^{-2/3}$ and 4.7438×10^{-16} $m^{-2/3}$, which fluctuated randomly; most of them were distributed between 3.41×10^{-18} $m^{-2/3}$ and 2.93×10^{-16} $m^{-2/3}$.

Table 32.4 Structure constants of optical signal samples

Measurement sample	Maximum ($m^{-2/3}$)	Minimum ($m^{-2/3}$)	Mean ($m^{-2/3}$)
T1	3.9475×10^{-21}	7.5496×10^{-22}	2.5071×10^{-21}
T2	4.6414×10^{-21}	1.1592×10^{-21}	3.0325×10^{-21}
T3	5.1554×10^{-21}	8.8539×10^{-22}	3.51×10^{-21}
T4	6.7502×10^{-21}	2.1776×10^{-21}	4.5620×10^{-21}
T5	1.0438×10^{-19}	4.9211×10^{-21}	4.9773×10^{-20}
T6	7.3957×10^{-20}	2.3715×10^{-21}	5.5249×10^{-20}
T7	1.2376×10^{-19}	3.9042×10^{-20}	7.5235×10^{-20}
T8	1.5142×10^{-19}	4.3730×10^{-20}	9.5812×10^{-20}
T15	2.3255×10^{-16}	5.5792×10^{-19}	1.0531×10^{-16}
T16	3.1901×10^{-16}	3.4070×10^{-18}	1.4859×10^{-16}
T18	4.7438×10^{-16}	2.6121×10^{-17}	1.8970×10^{-16}

According to the Davis standard, the turbulence intensity of the Chan and Ba rivers is weak and that of the Yantai No. 1 beach is moderate; the stronger the turbulence, the more unstable the signal. The fluctuation of the light signal sample in the Yantai No. 1 beach was the largest, followed by those of the Ba and Chan Rivers.

Probability Density Function

Light intensity in underwater turbulence fluctuates randomly. The probability density distribution is the most basic method for describing the statistical characteristics of actual underwater turbulence. At present, many studies have been conducted on the probability density distribution of light intensity fluctuation.

It is generally considered that the probability density function (PDF) of the light intensity fluctuation follows a log-normal distribution; its expression is [63]

$$f_I(I) = \frac{1}{I\sigma_{\ln I}\sqrt{2\pi}} \exp\left\{ -\frac{\left[\ln(I) + 0.5\sigma_{\ln I}^2\right]^2}{2\sigma_{\ln I}^2} \right\}, \tag{32.29}$$

where the logarithmic variance $\sigma_{\ln I}^2$ of the light intensity can be expressed by the scintillation index σ_I^2, in which I represents the light intensity, and $< \cdot >$ is the mean value in weak turbulence.

The gamma–gamma light intensity fluctuation probability distribution model is a two-parameter model that is often used to describe the statistics of light intensity fluctuation in the middle and strong fluctuation regions [64]. The light intensity at the receiving end can be regarded as a multiplicative modulation process of large-scale to small-scale light intensity fluctuation.

When both large- and small-scale light intensity fluctuations follow the gamma distribution and are statistically independent, the light intensity fluctuation follows the gamma–gamma distribution, where the probability density function is [64]

$$f_I(I) = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} I^{(\alpha+\beta)/2-1} K_{\alpha-\beta}\left(2\sqrt{\alpha\beta I}\right), \tag{32.30}$$

where I is the normalized irradiance, α and β are the outer and inner scale parameters used to describe the intensity scintillation, $\Gamma(\cdot)$ is the gamma function, and the second modified Bessel function is order $K_\nu(\cdot)$, when Gaussian beams are used. When the receiver aperture is used, the following are obtained:

$$\alpha = \left[\exp\left(\frac{0.196\sigma_l^2}{\left(1 + 0.18d^2 + 0.186\sigma_l^{12/5}\right)^{7/6}}\right) - 1\right]^{-1} \tag{32.31}$$

$$\beta = \left[\exp\left(\frac{0.204\sigma_l^2\left(1 + 0.23\sigma_l^{12/5}\right)^{-5/6}}{1 + 0.9d^2 + 0.207d^2\sigma_l^{12/5}}\right) - 1\right]^{-1}, \tag{32.32}$$

where $d = \sqrt{kD^2/4L}$ and $\sigma_l^2 = 1.23C_n^2k^{7/6}L^{11/6}$ is the Rytov variance under strong turbulence. $k = 2\pi/\lambda$, λ is the wavelength of light, D is the size of the receiving aperture, and C_n^2 is the refractive index structure constant.

Figure 32.12 shows the PDFs of the samples measured in different waters.

From Fig. 32.12a, b, we can see that the light intensity I on the receiving surface is normalized, and the value range of I increases with an increase in water velocity. It deviates to the left, and the distribution of $P(I)$ follows a log-normal distribution. It can be seen that an increase in water velocity will lead to an increase in turbulence intensity and an increase in the random fluctuation of the atmospheric refractive index, reducing the stability of the received optical signal.

As can be seen from Fig. 32.12c, d, the range of values of I in different water directions is basically the same, the probability density curves are basically coincidental, and the water direction has little effect on the change in the probability density function. It can be seen that the magnitude of the turbulence intensity is basically independent of the water direction when the laser propagates horizontally underwater.

Figure 32.12e shows the measurement results of the Yantai No. 1 beach at different distances. With an increase in the communication distance, the range of I also increases and it deviates to the left. Moreover, the combined turbulence

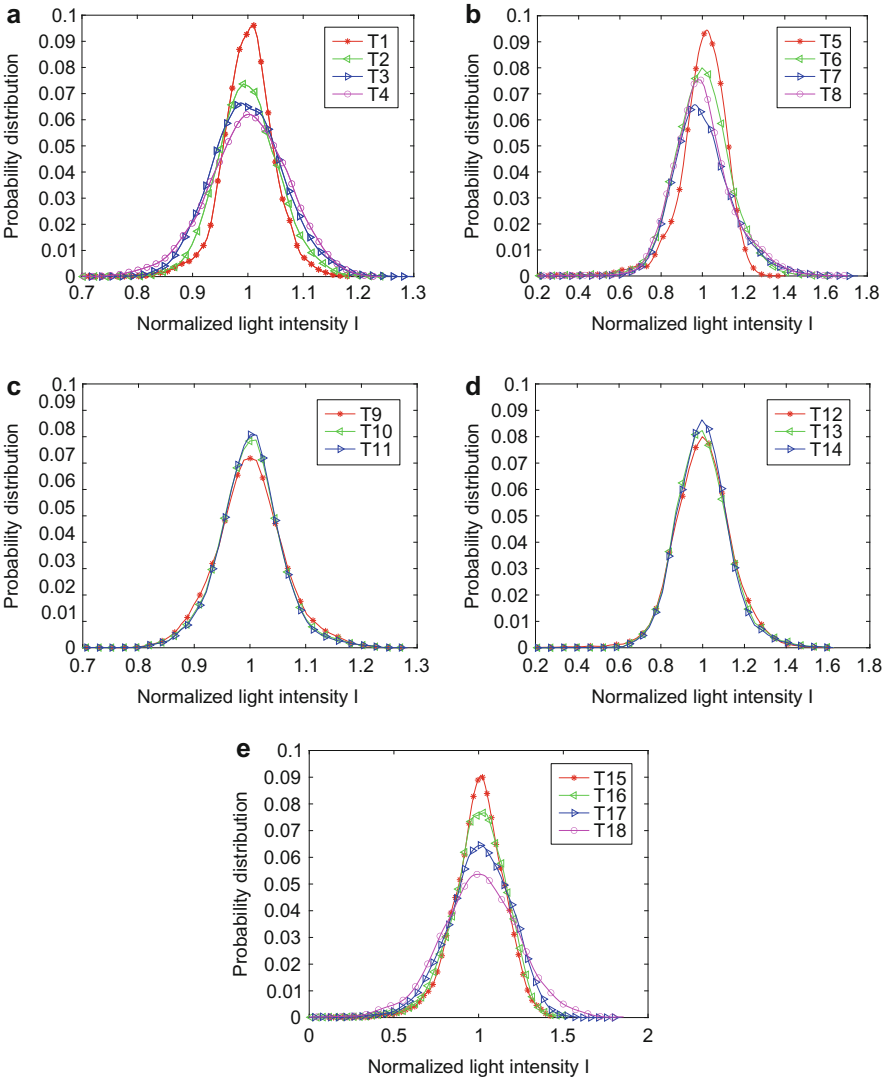


Fig. 32.12 PDFs of samples from different waters. (**a**) Chan River survey sample. (**b**) Ba River survey sample. (**c**) Chan River survey sample. (**d**) Ba River survey sample. (**e**) Samples taken at different distances from Yantai No. 1 beach

intensity is moderate, which is in accordance with the characteristics of the gamma–gamma distribution. Therefore, an increase in the distance increases the turbulence intensity, which reduces the stability of the optical signal received by the entire link.

Performance Analysis of Underwater Optical Wireless Communication Experiments

Underwater turbulence can affect the performance of communication systems, and even disrupt the system. In this section, the performance of an underwater communication system is further evaluated by studying the scintillation index, optical power loss, and bit error rate under different water speeds and distances.

Scintillation Index

The scintillation index is typically used to characterize the statistical characteristics of light intensity over time; the expression is [65]

$$\sigma_I^2 = \frac{\left(< I^2 > - < I >^2\right)}{< I >^2} = \frac{< I^2 >}{< I >^2} - 1, \tag{32.33}$$

where $< \cdot >$ denotes the ensemble average and I denotes the beam intensity. The normalized intensities in Fig. 32.13 were obtained by processing several groups of data from an underwater laser transmission experiment.

From Fig. 32.13, it can be seen that the normalized light intensity of the Chan River fluctuates randomly in the range of 0.9–1.1 that of the Ba River fluctuates randomly in the range of 0.8–1.2, and the intensity fluctuation range increases with the increase of water velocity. An increase in water velocity leads to an increase in the turbulence intensity and scintillation of the received optical signal, which reduces the stability of the received optical signal.

The normalized light intensity of Yantai No. 1 beach fluctuates randomly in the range of 0.5–1.5, and the fluctuation range increases with the communication distance. The intensity of underwater turbulence increases with an increase in the communication distance, and the scintillation of the received optical signal increases, reducing the stability of the received optical signal.

To explain quantitatively, the light intensity fluctuation at high water speeds is greater than that at low water speeds, and the fluctuation in light intensity at long distances is greater than that at short distances. Using Eq. (32.33), the scintillation index was calculated using the standard deviation of each set of samples.

The scintillation indices of the Chan River are 0.0032, 0.0041, 0.0050, and 0.0065, which are on the order of 10^{-3}. The scintillation indices of the Ba River are 0.045, 0.052, 0.060, and 0.068, which are on the order of 10^{-2}, from the quantitative scintillation index of the Chan and Ba Rivers. This proves that the fluctuation of light intensity at high water speeds is greater than that at low water speeds. The scintillation indices of the Yantai No. 1 beach are 0.27, 0.38, 0.52, and 0.61 at different distances, which are all on the order of 10^{-1}. This demonstrates that the turbulence intensity increases with an increase in the communication distance, by the quantitative scintillation index of the optical signal sample of the Yantai No. 1 beach.

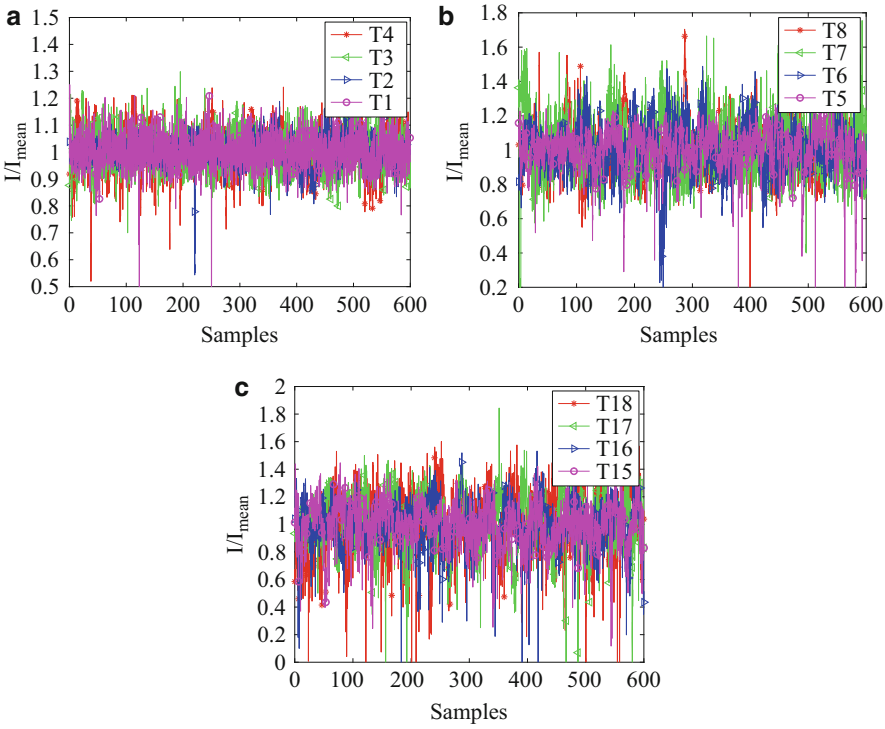


Fig. 32.13 Light intensity fluctuation of light signal samples in different waters. (**a**) Chan River. (**b**) Ba River. (**c**) Yantai No. 1 beach

Optical Power Loss

Based on Beer's theorem, the change in light power in seawater can be simplified as an exponential attenuation [66]:

$$I(L) = I_0 \exp(-c(\lambda)L), \tag{32.34}$$

where L is the distance of light propagation in water, I_0 is the initial light intensity, $I(L)$ is the received light intensity, and $c(\lambda)$ is the attenuation coefficient.

Figure 32.14 shows the power loss owing to turbulence in different waters. Figure 32.14a, b shows the decay curves of the received light power of the Chan and Ba Rivers, respectively, with the variation of the scintillation index at different water velocities. Figure 32.14c shows the decay curve of the received light power with the scintillation index at different distances of the Yantai No. 1 beach, and the loss of the light power increases with the scintillation index.

Comparing the measurements of the Chan River, Ba River, and Yantai No. 1 beach, we can see that the loss of the optical signal is different for different scintillation indices. The scintillation index of the Chan River was 10^{-3} orders of

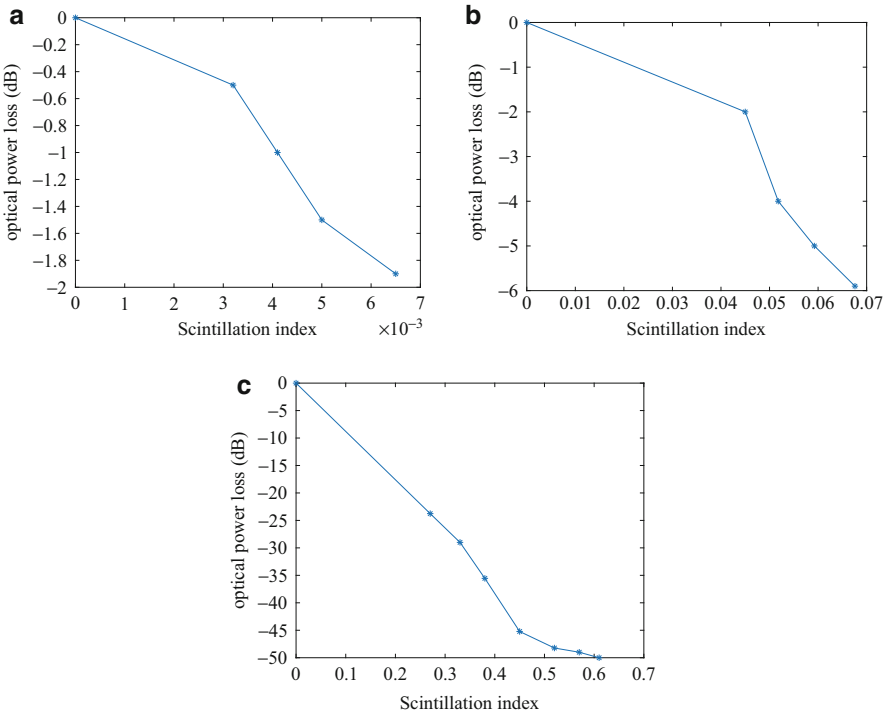


Fig. 32.14 Power loss due to turbulence in different water areas. (**a**) Power loss of the Chan River at different water speeds. (**b**) Power loss of the Ba River at different water speeds. (**c**) Power loss of the Yantai No. 1 beach at different distances

magnitude and the loss of the light signal was 1.9 dB; the Ba River's scintillation index was 10^{-2} orders of magnitude, and the loss of the light signal was 5.9 dB; the scintillation index of the Yantai No. 1 beach was 10^{-1} orders of magnitude at different communication distances, and the maximum loss of the optical signal was 50 dB.

Thus, the loss of the received optical signal increases with increasing water speed and communication distance. The Chan River, Ba River, and Yantai No. 1 beach can be transmitted for maximum distances of 1 m, 1.5 m, and 6.5 m, respectively. The corresponding attenuation coefficient of Yantai No. 1 beach is 2.13 and those of the Chan and Ba rivers are 10.79 and 9.21.

By comparing the longest distance that a laser could travel in different water qualities, it was found that the maximum distance was determined by the water attenuation coefficient, that is, the turbidity. Lasers can travel long distances in clear water and short distances in turbid water. As a result, the water quality of the Chan River was the worst, followed by that of the Ba River, and the water quality of the Yantai No. 1 beach was the best, which is consistent with reality.

BER at Different SNR

The instantaneous signal-to-noise ratio (SNR) probability density function $f_\mu(\mu)$ [67] of the received signal can be expressed as

$$f_\mu(\mu) = \frac{1}{2\mu\sigma_I\sqrt{2\pi}} \exp\left\{-\frac{\left[\ln(\mu/\overline{\mu}) + \sigma_I^2\right]^2}{8\sigma_I^2}\right\}. \tag{32.35}$$

In general, the average bit error rate in fading channels can be calculated by multiplying the instantaneous bit error rate by the probability density function $f_\mu(\mu)$ of the instantaneous signal-to-noise ratio of the received signal [68]:

$$\overline{P_e} = \int_0^\infty f_\mu(\mu) P_e d\mu, \tag{32.36}$$

where the instantaneous bit error rate $P_e = \frac{1}{2}\mathrm{erfc}\left(\frac{\sqrt{\mu}}{2}\right)$ is expressed by the average bit error rate formula [69]:

$$\overline{P_e} = \int_0^\infty \frac{1}{2\mu\sigma_I\sqrt{2\pi}} \exp\left[-\frac{\left(\ln(\mu/\overline{\mu}) + \sigma_I^2\right)^2}{8\sigma_I^2}\right] \times \frac{1}{2}\mathrm{erfc}\left(\frac{\sqrt{\mu}}{2}\right) d\mu, \tag{32.37}$$

where μ is the instantaneous signal-to-noise ratio, $\overline{\mu}$ is the average signal-to-noise ratio, and μ_{th} is the critical threshold of the SNR.

By introducing the probability density function of the gamma–gamma distribution into the above equation, we obtain [70]

$$\overline{P_e} = \frac{2^{\alpha+\beta-3}}{\pi^{3/2}\Gamma(\alpha)\Gamma(\beta)} \times G_{5,2}^{2,4}\left[\frac{4\overline{\mu}}{\alpha^2\beta^2}\middle|_{0,\frac{1}{2}}^{\frac{1-\alpha}{2},\frac{2-\alpha}{2},\frac{1-\beta}{2},\frac{2-\beta}{2},1}\right]. \tag{32.38}$$

Figure 32.15 shows the bit error rate for different signal-to-noise ratios in different water areas. Figure 32.15a, b show that the bit error rates of the four water velocities of the Chan River from small to large are 0.0133, 0.0134, 0.0135, and 0.0139, respectively; when the SNR is 10 dB, the bit error rates of the four water velocities in the Ba River from small to large are 0.0154, 0.0166, 0.0184, and 0.0210, respectively. When the signal-to-noise ratio is constant, the bit error rate of the sample with a high water speed is greater than that of the sample with a low water speed.

In Fig. 32.15c, when the signal-to-noise ratio is 10 dB, the corresponding bit error rates of the Yantai No. 1 beach are 0.0297, 0.0344, 0.0390, and 0.0449 when the communication distances are 1 m, 3 m, 5 m, and 6.5 m, respectively. The greater the communication distance, the higher the bit error rate. When the bit error rate is 10^{-2}, the SNRs of the Chan River are 11.0, 11.05, 11.1, and 11.15 and those of the Ba River are 13.5, 14, 14.5, and 15, respectively. The higher the water speed, the higher

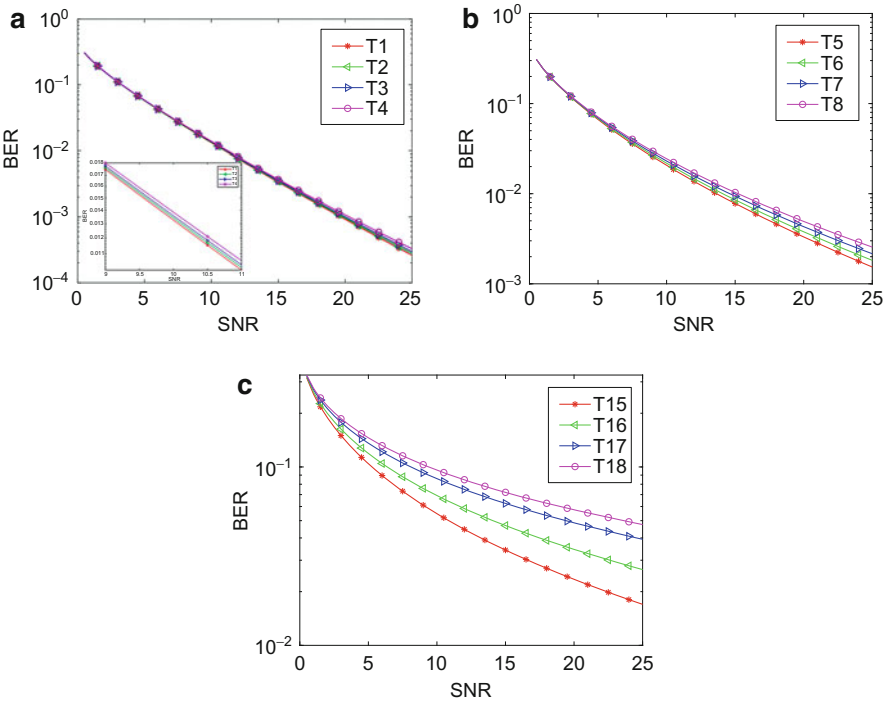


Fig. 32.15 Bit error rates at different signal-to-noise ratios for different water areas. (**a**) Chan River. (**b**) Ba River. (**c**) Yantai No. 1 beach

the signal-to-noise ratio. When the bit error rate is 0.15, the signal-to-noise ratios corresponding to different distances from the Yantai No. 1 beach are 5, 6, 7, and 8, respectively.

Figure 32.15 shows that the difference in bit error rate caused by different water speeds in the Ba River was greater than that caused by the Chan River. The change of the bit error rate of the Chan River at different water speeds was smaller than that of the Ba River at different water speeds.

Optical Cross-Medium Transmission Characteristics

At the air–sea interface, the laser propagates from a sparse medium to a dense medium. The random fluctuation of the sea surface leads to the spread, flickering, and drift of the receiving surface spot, which changes the communication performance. Underwater cross-medium optical communication is affected not only by the absorption and scattering of the light spot centroid but also by the sea surface wind speed. The spatial intensity distribution of the spot and offset of the spot centroid propagating across the medium were analyzed.

Optical Transmedia Model

Owing to the influence of the sea surface wind speed, the sea surface fluctuates randomly over time. When the laser irradiates the sea surface, the different refractive indices of the atmosphere and seawater cause the transmission laser to refract and reflect. Figure 32.16 shows the refraction of light at the air–sea interface, assuming that the refractive index of the medium in which the incident light resides is n_1 and the refractive index of the medium in which the refracted light resides is n_2.

The relationship between the incident angle θ_i, reflection angle θ_r, and refraction angle θ_t in the local sea surface coordinate system satisfies the reflection and Snell's laws:

$$\theta_i = \theta_r. \tag{32.39}$$

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{n_1}{n_2}. \tag{32.40}$$

The proportion R of the power of the incident light reflected by the air–sea interface to the power of the incident light is called reflectivity; the reflectivity of vertically polarized light is R_s, and the reflectivity of parallel-polarized light is R_p. Then,

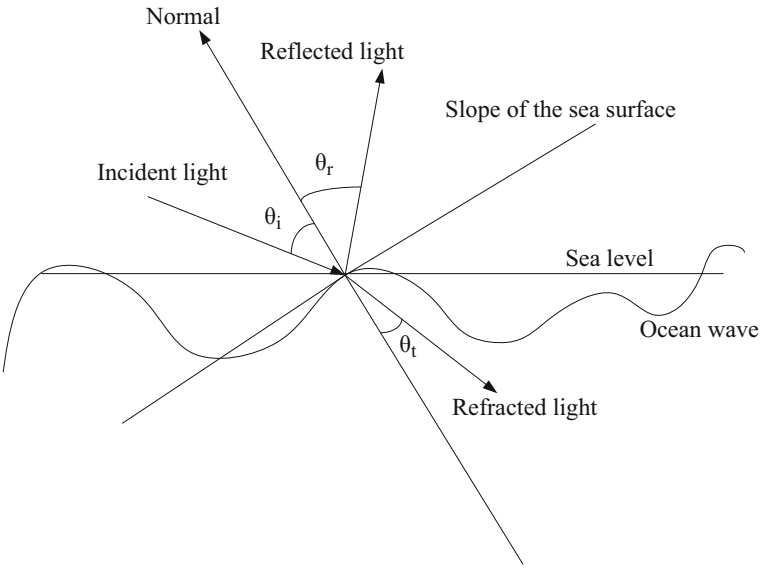


Fig. 32.16 Simplified model of the air–sea interface for light transmission

$$R_s = \frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)} \tag{32.41}$$

$$R_p = \frac{\tan^2(\theta_i - \theta_t)}{\tan^2(\theta_i + \theta_t)}. \tag{32.42}$$

The ratio of the power of the refracted light to that of the incident light is called the transmittance, T:

$$R = \frac{1}{2}\left(R_s + R_p\right) \tag{32.43}$$

$$T = 1 - R. \tag{32.44}$$

Considering the cross-medium light transmission background, a model of the center shift of the light spot caused by underwater cross-medium turbulence using a Gaussian function was proposed. The probability density function is defined as follows:

$$f(I) = A \times \frac{1}{\sqrt{2\pi} \times \sigma} \times e^{\left|-\frac{(x-\mu)^2}{2\times\sigma^2}\right|}, \tag{32.45}$$

where A is the normalized correction coefficient; the values are in the range of 0–1. $\int_0^\infty f(I) = 1$; hence, μ is the mean value and σ is the dithering variance of the spot center offset in the current turbulent state.

The Gaussian function was used to fit the cumulative distribution function of the spot center offset caused by the cross-medium turbulence:

$$F(I) = \int_0^I f(x)dx = \int_0^I A \times \frac{1}{\sqrt{2\pi} \times \sigma} \times e^{\left|-\frac{(x-\mu)^2}{2\times\sigma^2}\right|}dx. \tag{32.46}$$

When the signal-to-noise ratio of the optical communication system is less than γ_{th}, the communication system is interrupted:

$$P_{\text{out}} = F(\gamma_{\text{th}}) = \int_0^{\gamma_{\text{th}}} A \times \frac{1}{\sqrt{2\pi} \times \sigma} \times e^{\left|-\frac{(x-\mu)^2}{2\times\sigma^2}\right|}dx. \tag{32.47}$$

Therefore, the general expression of the bit error rate for the different binary modulation modes is

$$P_{\text{BER}} = \frac{q^p}{2\Gamma(p)} \int_0^\infty \exp(-qx)x^{p-1}F(x)dx, \tag{32.48}$$

where the different values of parameters p and q represent different modulation methods. $P = 1$ and $q = 1$ represent differential phase-shift keying.

Experiment

Spatial Distribution of the Spot Intensity

To obtain the spatial distribution of the light intensity in different water areas, several cross-medium laser transmissions were carried out in the Chan and Ba rivers in Xi'an City, Shaanxi Province, and in the upper reaches of Qianyou River in Shangluo City, Zhashui County, Shaanxi Province. In the experiment, an electronic thermometer was used to measure the temperature; the floating method was used to measure the water velocity many times to obtain the average value. Some of the experimental results are listed in Table 32.5.

To obtain the spatial distribution of the light intensity in different water areas, the image of the light spot collected by a charge-coupled device (CCD) camera was normalized. Typical images of light spots are shown in Fig. 32.17.

As shown in Fig. 32.17, the intensity distribution of the underwater wireless light across the medium follows a Gaussian distribution. The spatial distribution of the light intensity changes with the water quality. It can be seen from the picture that the light spot image of the Qianyou River is more concentrated and the light spot image of the Chan River is more diffuse. The distortion and dispersion degree of the light spot can be divided into the Chan, Ba, and Qianyou rivers, and the water quality of the Chan, Ba, and Qianyou rivers ranges from poor to good. The worse the water quality, the greater the degree of absorption and scattering and the greater the degree of distortion and dispersion of the light spot.

To analyze the difference in light intensity distribution in different waters, a two-dimensional distribution of the light intensity was obtained by projecting a gray image along the x-/y-axis. The results are shown in Fig. 32.18.

As shown in Fig. 32.18, the spatial distribution of the light spot intensity fluctuates randomly, and the shape of the light spot and degree of energy attenuation differ in different waters. The steepness of the light spot intensity distribution obtained from the Qianyou River was the highest, followed by that from the Ba River, and the Chan River was the smoothest.

With an increase in the attenuation coefficient, the water quality worsens, and the number of absorbed and scattered photons in seawater increases. Because of the increase in water turbidity, the absorption and scattering of photons were enhanced, and the intensity of the light spot was weakened.

The maximum normalized gray values of the received light spot in the Qianyou, Ba, and Chan rivers were 0.2550, 0.2318, and 0.2125, respectively. The light

Table 32.5 Environmental conditions of the experiments

Water	Date	Water temperature (° C)	Wind speed (m/s)	Water speed (m/s)
Chan River	2023.5.9	20–21.5	1	0.45
Ba River	2023.5.13	21.3–22.9	1	0.47
Qianyou River	2023.6.10	19.1–20.3	2	0.5

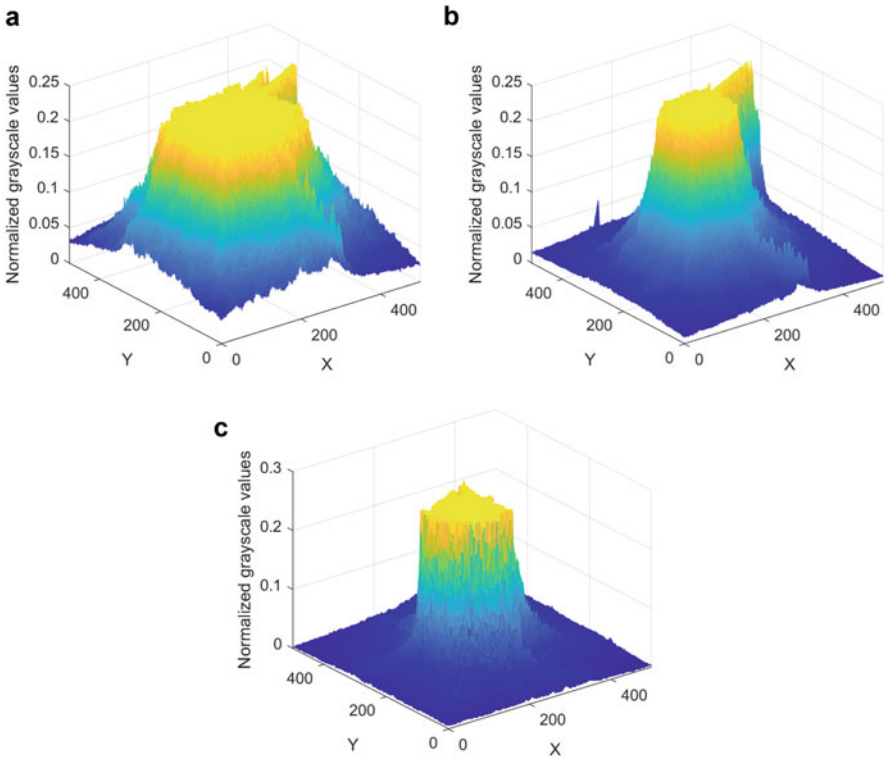


Fig. 32.17 Three-dimensional image of the light intensity in different waters. (**a**) Chan River. (**b**) Ba River. (**c**) Qianyou River

intensity of the Qianyou River was the lowest, followed by those of the Ba and Chan Rivers. In the light intensity distribution curve, we can observe the alternating appearance of the intensity distribution, which is caused by the diffraction phenomenon in the laser transmission process, resulting in light and dark rings on the receiving surface.

Spot Center Offset

The probability density function of the centroid offset of the light spot under different water conditions can be obtained by drawing the centroids of many light spot images measured over a long time under different water conditions, as shown in Fig. 32.19.

Figure 32.19 shows the probability distributions of the centroid offsets of light spots measured in the upper reaches of the Chan, Ba, and Qianyou rivers, respectively, which all satisfy the Gaussian distribution. From Fig. 32.19, it can be seen that the normalized offsets of the light spot centroid measured at the Chan, Ba, and Qianyou rivers are between 0.96–1.04, 0.92–1.08, and 0.85–1.15, respectively.

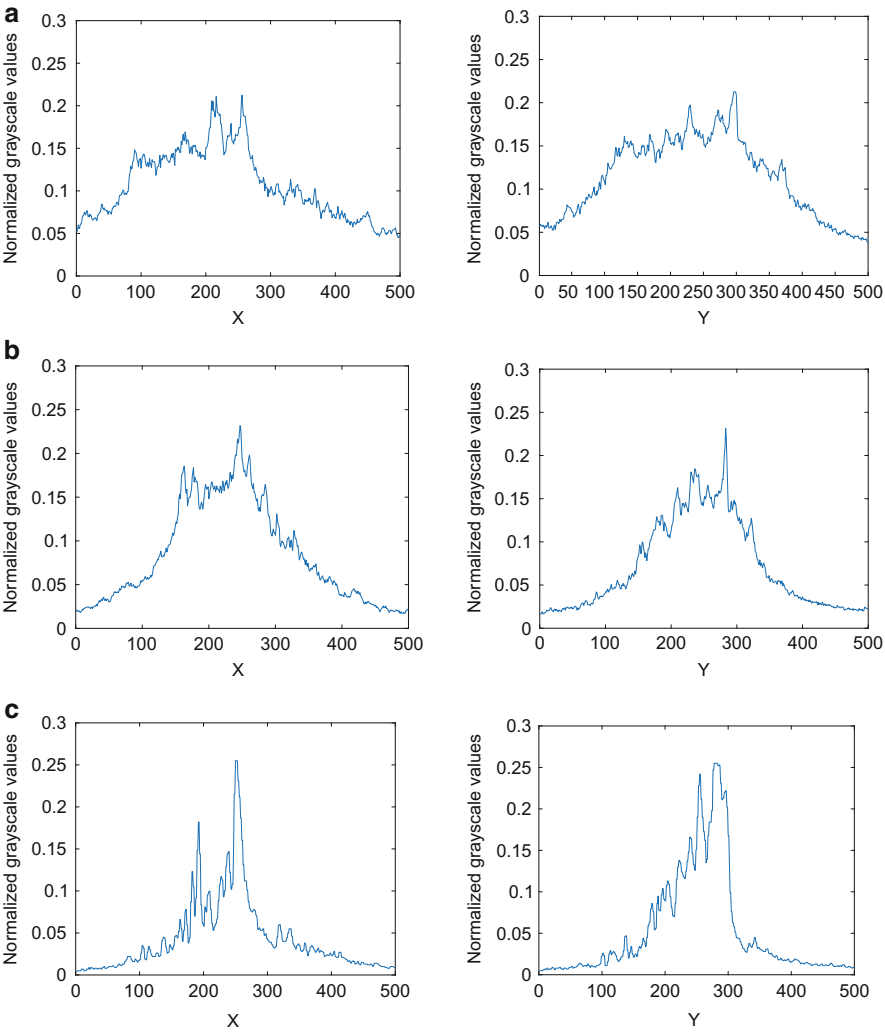


Fig. 32.18 Two-dimensional projection of the light intensity over different waters. (**a**) Chan River. (**b**) Ba River. (**c**) Qianyou River

With an increase in wind speed on the water surface, the random fluctuation of the water surface becomes more severe, which leads to an increase in the probability of a large sea surface inclination angle. When the photons project out of the sea surface, the direction deviates farther from the center of the receiving surface. In addition to the limitation of the receiving field angle, the photons sometimes fall outside the receiving field angle; thus, the number of photons on the receiving surface decreases and they are not included in the statistical results. As a result, the weight of the center of the receiving surface gradually decreases.

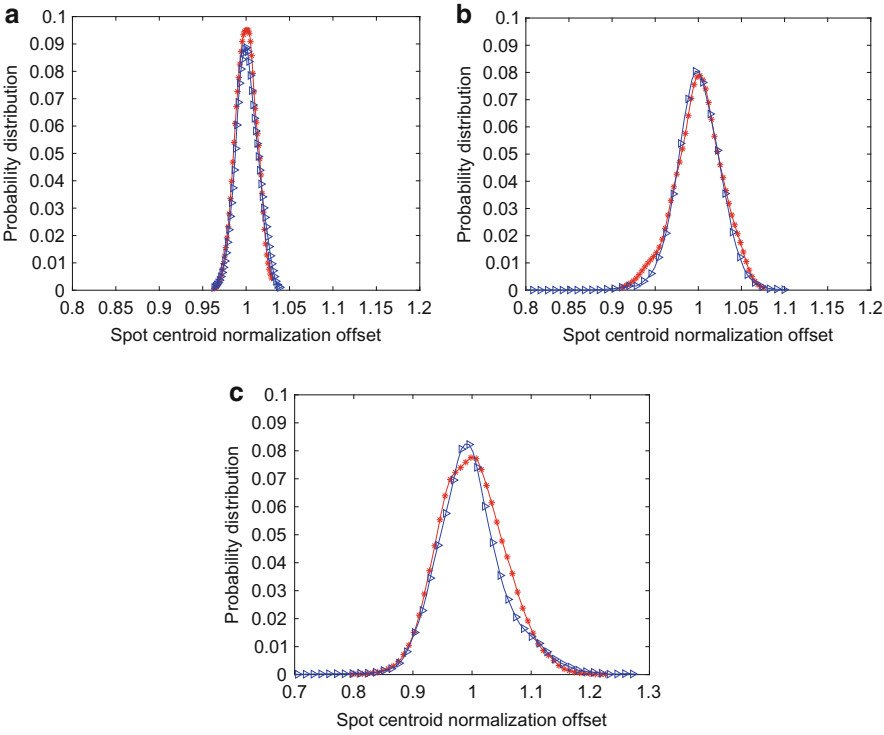


Fig. 32.19 Probability density distribution of the light spot centroid offset in different waters. (**a**) Chan River. (**b**) Ba River. (**c**) Qianyou River

The following experimental results can be obtained:

1. The spatial distribution of the light spot intensity follows the Gaussian distribution in different water areas, and the distortion and dispersion of the light spot image are higher in areas with poor water quality.
2. The light spot centroid offset of the cross-medium optical communication conforms to the Gaussian model.
3. The probability density distribution of the light spot centroid offset differs in different water areas. Under the experimental conditions, the light spot centroid offset is larger in water areas with higher wind speeds, higher water speeds, and severe water surface fluctuation.

References

1. Duntley S Q. Light in the sea [J]. Journal of the Optical Society of America, 1963, 53(2): 214–233.
2. Gilbert G D, Stoner T, Jernigan J L. Underwater experiments on the polarization, coherence, and scattering properties of a pulsed blue-green laser [J]. Proceedings of SPIE-The International Society for Optical Engineering, 1966, 12(1):7–8.

3. Petzold T J. Volume scattering functions for selected ocean waters [R]. Scripps Institution of Oceanography La Jolla Ca Visibility Lab, 1972.
4. Jaruwatanadilok S. Underwater wireless optical communication channel modeling and performance evaluation using vector radiative transfer theory [J]. IEEE Journal on Selected Areas in Communications, 2008, 26(9): 1620–1627.
5. Cochenour B M, Mullen L J, Laux A E. Characterization of the beam-spread function for underwater optical wireless communications links [J]. IEEE Journal of Oceanic Engineering, 2009, 33(4):513–521.
6. Cox, Willian C, Jr A, et al. Simulation, modeling, and design of underwater optical communication systems [M]. North Carolina State University, 2012.
7. Li Jing. Research on carrier modulated underwater laser communication based on analytical Monte Carlo method [D]. Wuhan: Huazhong University of Science and Technology, 2013.
8. Nikishov V V, Nikishov V I. Spectrum of turbulent fluctuations of the sea-water refraction index [J]. International Journal of Fluid Mechanics Research, 2000, 27(1):82–98.
9. Lu Lu. Effect of ocean turbulence on laser beam propagation [D]. Hefei: University of Science and Technology of China, 2016.
10. Huang Y, Zhang B, Gao Z, et al. Evolution behavior of Gaussian Schell-model vortex beams propagating through oceanic turbulence [J]. Optics Express, 2014, 22(15): 17723–17734.
11. Lu L, Ji X, Li X, et al. Influence of oceanic turbulence on propagation characteristics of gaussian array beams [J]. Optik - International Journal for Light and Electron Optics, 2014, 125(24): 458–476.
12. Mcglamery B L. Restoration of turbulence-degraded images [J]. Journal of the Optical Society of America, 1967, 57(3):293–296.
13. Roddier N A. Atmospheric wavefront simulation using zernike polynomials [J]. Optical Engineering, 1990, 29(10):1174–1180.
14. Cui Z, Yue P, Xiang Y, et al. Effect of convergent beam array on reducing scintillation in underwater optical wireless communications with pointing errors [J]. Optics Express, 2021, 29(7): 9846–9860.
15. Zhang J, Kou L, Yang Y, et al. Monte-Carlo-based optical wireless underwater channel modeling with oceanic turbulence [J]. Optics Communications, 2020, 475(1): 126214.
16. Yi X, Li Z, Liu Z. Underwater optical communication performance for laser beam propagation Through weak oceanic turbulence. [J]. Appl Optics, 2015,54(6) : 1273–8.
17. Fu Y, Du Y. Performance of heterodyne differential phase-shift-keying underwater optical wireless communication systems in gamma-gamma-distributed turbulence [J]. Appl Optics, 2018,57(9):2057–2063.
18. Oubei H M, Zedini E, ElAfandy R T, et al. Efficient Weibull channel model for salinity induced turbulent underwater optical wireless communications [C] Opto-Electronics and Communications Conference and Photonics Global Conference. IEEE, 2017: 1–2.
19. Jamali M V, Khorramshahi P, Tashakori A, et al. Statistical distribution of intensity fluctuations for underwater wireless optical channels in the presence of air bubbles [C] Iran Workshop on Communication and Information Theory. IEEE, 2016: 1–6.
20. Zedini E, Oubei H M, Kammoun A, et al. A new simple model for underwater wireless optical channels in the presence of air bubbles [C] Global Communications Conference. IEEE, 2017: 1–6.
21. Duntley S Q. Light in the sea [J]. JOSA, 1963,53(2):214-233.
22. Doherty J J. Submarine communication [J]. The Journal of the Acoustical Society of America, 1981, 69(1):335–335.
23. Puschell J J, Giannaris R J, Stotts L. The Autonomous Data Optical Relay Experiment: first two way laser communication between an aircraft and submarine [C], [Proceedings] NTC-92: National Telesystems Conference. IEEE, Washington, DC, United States, 1992: 14/27–14/30.
24. Tivey M, Fucile P, Sichel E. A low power, low cost, underwater optical communication system [J]. Ridge Events, 2000, 1(01): 27–29.

25. Raj B D, Rakesh B S, Avinash N, et al. Visible spectrum using optical communication [J]. Perspectives in Communication, Embedded-systems and Signal-processing-PiCES, 2018, 2(7): 153–156.
26. Hanson F, Radic S. High bandwidth underwater optical communication [J]. Applied optics, 2008, 47(2):277–283.
27. Cossu G, Corsini R, Khalid A M, et al. Experimental demonstration of high speed underwater visible light communications[C], Optical Wireless Communications (IWOW), 2013 2nd International Workshop on. IEEE, Newcastle Upon Tyne, UK, 2013:11–15.
28. Nakamura K, Mizukoshi I, Hanawa M. Optical wireless transmission of 405 nm, 1.45 Gbit/s optical IM/DD-OFDM signals through a 4.8 m underwater channel [J]. Optics express, 2015, 23(2):1558–1566.
29. Oubei H M, Durán J R, Janjua B, et al. Wireless optical transmission of 450 nm, 3.2 Gbit/s 16-QAM-OFDM signals over 6.6 m underwater channel [C], Lasers and Electro-Optics (CLEO), 2016 Conference on. IEEE, California, United States, 2016:1–2.
30. Tsai W S, Lu H H, Wu H W, et al. A 30 Gb/s PAM4 underwater wireless laser transmission system with optical beam reducer/expander [J]. Scientific reports, 2019, 9(1):1–8.
31. Kong M, Guo Y, Alkhazragi O, et al. Real-time Optical-Wireless Video Surveillance System for High Visual-fidelity Underwater Monitoring [J]. IEEE Photonics Journal, 2022, 14(2):1–11.
32. Zhou Tianhua, He Ning, Ao Faliang. DSP implementation of underwater optical communication PPM digital receiver [J]. Optics, 2006, 32(s 1): 607–609.
33. Luo Linfeng, Wang Shuo. Design of optical communication and ranging in underwater mobile sensor networks [J]. Robotics and applications, 2008, 02(02): 36–39.
34. Sui Meihong, Yu Xinsheng, Liu Xifeng, Zhou Zhangguo. Study on the characteristics of seawater channel in underwater optical wireless communications [J]. Ocean Sciences, 2009, 33(06): 80–85.
35. Charlotte. Design and implementation of high-speed underwater optical communication system based on LED [D]. Qingdao, Shandong: Ocean University of China, 2013:1–3.
36. Shang Tao, Du Gang, Zhang Xiaoyu, Zhang Hongkun. Underwater visible light communication node design [J]. Journal of Huazhong University of Science and Technology (natural science edition), 2015, 43(s 1): 211–215.
37. Xu J, Kong M, Lin A, et al. OFDM-based broadband underwater optical wireless communication system using a compact blue LED [J]. Optics Communications, 2016, 369(01):100–105.
38. Huang Y F, Chi Y C, et al. Filtered multicarrier OFDM encoding on blue laser diode for 14.8-Gbps seawater transmission [J]. Journal of Lightwave Technology, 2017, 36(9): 1739–1745.
39. Hong X, Fei C, Zhang G, et al. Discrete multitone transmission for underwater optical wireless communication system using probabilistic constellation shaping to approach channel capacity limit [J]. Optics Letters, 2019, 44(3):558–561.
40. Wang J, Lu C, Li S, et al. 100 m/500 Mbps underwater optical wireless communication using an NRZ-OOK modulated 520 nm laser diode [J]. Optics Express, 2019, 27(9): 12171–12181.
41. Lu Weichao, Xu Jing. Underwater optical wireless communication system based on spread spectrum technology [J]. Radio communications technology, 2021, 47(04): 410–415.
42. Sun Y. Experimental Study on Underwater Blue Green Optical Communication System [D]. Xi'an University of Technology, 2022.
43. Kaushal H, Kaddoum G. Underwater Optical Wireless Communication IEEE Access, 2016, 4(1): 1518–1547.
44. Smith R C, Baker K S. Optical properties of the clearest natural waters [J]. Applied Optics, 1981, 20(2):177–184.
45. Morel A. Optical properties of pure water and pure sea water [M]. New York: Academic Press, 1974, 1–24.
46. Morel A, Maritorena S. Bio-optical properties of oceanic waters: A reappraisal [J]. Journal of Geophysical Research: Oceans, 2001, 106(4):7163–7180.

47. Prieur L, Sathyendranath S. An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials 1 [J]. Limnology and Oceanography, 1981, 26(4): 671–689.
48. Morel A, Prieur L. Analysis of variations in ocean color 1 [J]. Limnology and oceanography, 1977, 22(4):709–722.
49. Wu Yongsen, Zhang Shikui, Zhang Xuqin, et al. . Experimental study on light absorption properties of sea water yellow matter [J] . Oceans and lakes, 2002, 33(04): 402–406.
50. Bricaud A, Morel A, Prieur L. Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains [J]. Limnol. Oceanogr, 1981, 26(1):43–53.
51. Pegau W S, Cleveland J S, Doss W, et al. A comparison of methods for the measurement of the absorption coefficient in natural waters [J]. Journal of Geophysical Research: Oceans, 1995, 100(7):13201–13220.
52. Gilerson A, Zhou J, Hlaing S, et al. Fluorescence component in the reflectance spectra from coastal waters. Dependence on water composition [J]. Optics Express, 2007, 15(24): 15702–15721.
53. Haltrin V I. Chlorophyll-based model of seawater optical properties [J]. Applied Optics, 1999, 38(33):6826–6832.
54. Roesler C S, Boss E. Spectral beam attenuation coefficient retrieved from ocean color inversion [J]. Geophysical Research Letters, 2003,30(9):1468–1469.
55. Poulin C, Zhang X, Yang P, et al. Diel variations of the attenuation, backscattering and absorption coefficients of four phytoplankton species and comparison with spherical, coated spherical and hexahedral particle optical models [J]. Journal of Quantitative Spectroscopy and Radiative Transfer, 2018, 217(01):288–304.
56. Stramski D, Bricaud A, Morel A. Modeling the inherent optical properties of the ocean based on the detailed composition of the planktonic community [J]. Applied Optics, 2001, 40(18): 2929–2945.
57. Ke X Z, Yang S J, Sun Y, et al. Underwater blue-green LED communication using a double-layered, curved compound-eye optical system [J]. Optical Express. 2022, 30(11):18599–18616.
58. Li Jun, Luo Jianghua, Yuan Xiuhua. Effect of water surface wave disturbance on optical wireless communication [J]. Journal of Optics, 2021, 41(07): 47–54.
59. Jiang Yu, Chen Xun, Tao Junyong, et al. . Numerical simulation of non-gaussian stochastic processes with specified spectral density, skew and kurtosis values [j]. Journal of systems simulation, 2006, 18(5): 1127–1130.
60. Yahya Baykal. Expressing oceanic turbulence parameters by atmospheric turbulence structure constant [J]. Applied Optics. 2016, 55(6):1228–1231.
61. Arnold T. Optical turbulence parameters characterized via optical measurements over a 2.33 km free-space laser path [J]. Optical Express, 2008, 16(19):14645–14654.
62. Wu Xiaojun, Wang Hongxing, song Bo, etc.. Measurement and transmission characteristics of light intensity fluctuations in different environments [J]. Photoelectron Laser, 2015, 26(6): 1138–1145.
63. Pang W N, Wang P, Guo L X, et al. Performance investigation of UWOC system with multiuser diversity scheduling schemes in oceanic turbulence channels [J]. Optics Communications, 2019, 441(1):138–148.
64. he fengtao, Du Ying, Zhang Jianlei, et al.. Bit error rate of pulse position modulation optical wireless communication in Gamma-Gamma ocean anisotropic turbulence [J]. Journal of Physics, 2019, 68(16): 236–244.
65. Ji Xiuyang, Yin Hongxi, Jing Lianyou, etc. . Performance analysis of underwater optical wireless communication system with strong turbulence channel based on strong wave theory [J]. Journal of Optics, 2022,42(18): 9-16.
66. Huang Aiping, Zhang Yingluo, Tao Linwei. Monte Carlo simulation of underwater laser communication channel characteristics [J] . Infrared and laser engineering, 2017, 46(04): 226–231.

67. Zhu X M, Kahn J M. Free-space optical communications through atmospheric turbulence channels [J]. IEEE Transactions on Communications, 2002, 50(8): 1293–1300.
68. Chen Quanrun, Yu Xiang, Cui Wennan, et al.. Visible Light Communication and performance analysis based on short-and medium-range inter-satellite links [J]. Journal of optics, 2019, 39(10): 102–112.
69. Qin Yunyi, Xie mengtong, Xu Zhiyong, etc.. Performance analysis of free space optical communication system based on airy light [J]. Optical communications technology, 2022, 46(04): 46–50.
70. Fu Yuqing, Duan Qi, Zhou Lin. Study on the performance of underwater optical wireless communication system with Gamma-Gamma strong ocean turbulence and aiming error [j]. Infrared and laser engineering, 2020, 49(02): 110–117.

Principles and Research Progress on LEDs 33

Contents

Introduction 1258
Basic LED Principles and Photoelectric Characteristics 1258
LED Fundamentals 1258
Electrical Properties of LEDs 1262
Optical Properties of LEDs 1264
LED Research and Status 1266
Research and Development of Red LEDs 1266
Research and Development of Yellow LEDs 1268
Research and Development of Green LEDs 1269
Research and Development of Blue LEDs 1273
Research and Development of White LEDs 1277
Research and Development of UV LEDs 1280
LED Device Material History 1282
Visible Light Full-Duplex Communication 1284
Research and Development of Visible Light Communication 1284
Overall Full-Duplex Communication System Structure 1286
LED Optical Design 1286
LED Driver Circuit 1289
Experimental Verification 1292
Overall Uplink Design 1292
System Experiment 1293
Summary 1295
LED Industry Opportunities 1296
LED Industry Challenges 1297
References 1298

Abstract

LED is highly valued by the society for its energy-saving, efficient, durable, green, environmental protection, and other advantages and has a wide range of applications in display screens, traffic lights, LCD screen backlights, greenhouse fill lights and daily lighting, and even communications. Starting from the principle of LED, this paper introduces the optical and electrical characteristics of LED,

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_33

discusses the research and progress of red, green, yellow, and blue LEDs, and expounds on the light-emitting principle and research progress of white LED. The application of white LED in visible full-duplex communication is summarized, and finally, the future development of LED is prospected.

Keywords

LED · White LED · Visible light communication · Optical design · Channel modeling

Introduction

With the rapid development of science and technology, incandescent lamps, and fluorescent lamps, traditional light sources with low efficiency and high energy consumption are gradually moving into the historical stage [1]. Light-emitting diodes (LEDs) have attracted attention in the market, owing to their advantages of energy savings, long life cycle, high economic benefit, stable performance, environment-friendliness, and low-cost performance [2].

At present, LEDs are widely used in outdoor solid-state lighting equipment, such as street lamps and screen displays [3]. With the rapid development of LED lighting technology, LED energy efficiency has reached a very high level [4]; at the same lighting efficiency, LED lamp energy consumption is one-eighth that of incandescent lamps and half that of energy-saving lamps. Hence, they are more in line with the energy-saving green development concept. It is believed that, in the near future, LED lighting products will completely replace traditional lighting products [5].

At present, LEDs have entered the commercial application stage and become mainstream lighting, providing great convenience to people's lives. According to a 2021 report on the market prospects and investment opportunities of China's LED lighting industry, by the end of 2020, the global penetration rate of LED lighting had reached 59.0%, while the domestic penetration rate had reached 78%. This leaves room for more development in the future [6].

Basic LED Principles and Photoelectric Characteristics

LED Fundamentals

An LED is a widely used semiconductor light source. It uses the carrier radiation of the P–N junction composite luminescence, which can be used for optical communication. Typical luminescence wavelengths are 780 nm, 850 nm, 1300 nm, etc. LEDs are inexpensive and reliable, the driving circuit is simple, the characteristics are only slightly affected by the temperature, and the transmitting power and driving current have a good linear relationship.

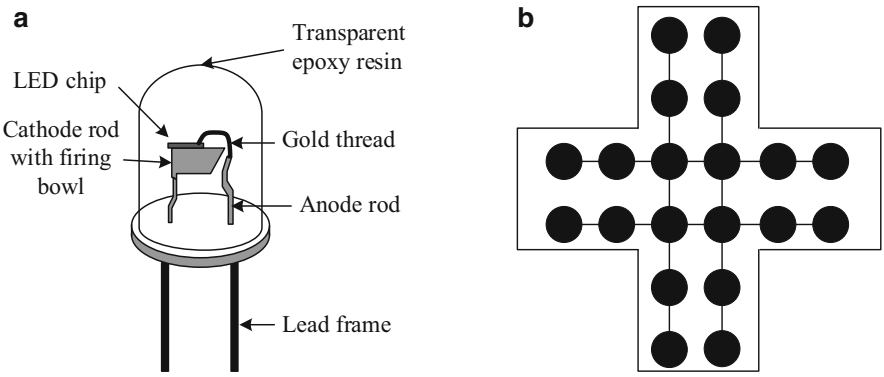


Fig. 33.1 Basic framework and LED surface light source. (**a**) Basic LED device structure [7] (**b**) LED surface light source [8]

The basic structure of common LED devices is shown in Fig. 33.1a [7]. Such LEDs are typical point sources with a high directivity. If multiple LED chips are packaged on a panel, a surface light source is formed that still has high directivity, as shown in Fig. 33.1b [8].

Similar to other diodes, the core of a light-emitting diode is a semiconductor chip consisting mainly of a P–N junction, which is unidirectionally conductive, as shown in Fig. 33.2 [9]. The P–N junction consists of two semiconductors: a p-type and an n-type, whose main carriers are holes and electrons, respectively. When a forward current passes through, the holes and electron carriers in the P–N junction flow to the N and P regions, respectively. The holes are recombined with the electrons, and the excess energy is emitted as photons [10].

If the energy emitted is in the range of visible light (390–760 nm), it can be seen by the human eye. The color depends on the size of the semiconductor band gap. In general, the smaller the band gap in the semiconductor materials, the less energy is emitted, and the longer the wavelength of the emitted light. Conversely, the larger the band gap, the greater the energy emitted by the photons and the shorter the wavelength of the emitted light.

For example, indium gallium nitride (InGaN) can be used as a blue light LED material, owing to its large band gap [11], while aluminum gallium phosphide (AlGaP) has a slightly smaller band gap and can be used as a green light LED material [12]. Aluminum gallium arsenide (AlGaAs) can be used as a red LED material, owing to its smaller band gap [13].

There are three main mechanisms for recombining excess carriers in a double heterojunction, depending on the energy exchange mode: Shockley–Read–Hall (SRH) recombination, radiative recombination, and Auger recombination [14]. Figure 33.3 shows a schematic representation of these three composite mechanisms, where (a) is the SRH recombination, (b) is the radiative recombination, and (c) is the Auger recombination.

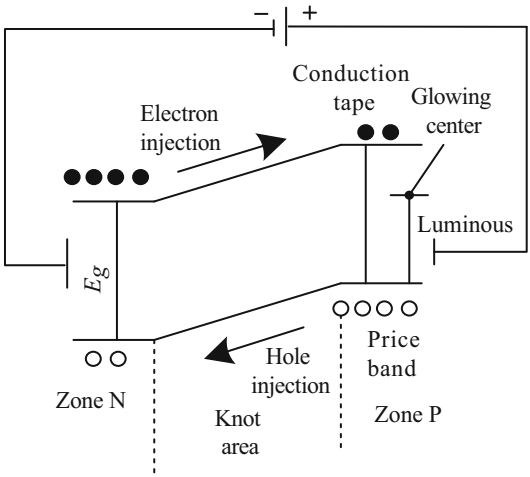


Fig. 33.2 LED emission principle

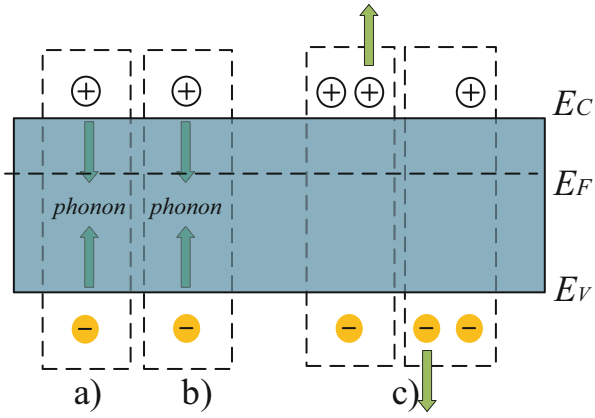


Fig. 33.3 Schematic diagram of three types of electron and hole recombination

SRH Recombination

SRH recombination is related to the existence of traps or defect energy levels, which is mainly the deep-level recombination of free carriers. The SRH recombination consists of two processes: a neutron process in which the defect center captures an electron to release a neutron and a phonon process in which the defect center loses an electron to release a phonon. The carrier-transport process can be expressed as [15]

$$\mathrm{Trap}^0 + e^- \rightarrow \mathrm{Trap}^- + E_{p1} \tag{33.1}$$

$$\mathrm{Trap}^- + h^+ \rightarrow \mathrm{Trap}^0 + E_{p2}. \tag{33.2}$$

$Trap^0$ and $Trap^-$ represent neutral and electronegative defect centers, e^- represents electrons, and h^+ represents holes. While E_{p1} represents the bandgap width between the conduction band E_C and the defect level E_T, E_{p2} represents the bandgap width between the defect level E_T and the valence band E_V:

$$E_{p1} = E_C - E_T \tag{33.3}$$

$$E_{p2} = E_T - E_V. \tag{33.4}$$

The non-radiative recombination rate of the defect center with a defect energy level of E_T is

$$R_{SRH} = \frac{p_0\Delta n + n_0\Delta p + \Delta n\Delta p}{\left(N_T v_p \sigma_p\right)^{-1}(n_0 + n_1 + \Delta n) + \left(N_T v_n \sigma_n\right)^{-1}(p_0 + p_1 + \Delta p)}, \tag{33.5}$$

where P_0 and N_0 are the equilibrium concentrations of holes and electrons, respectively; p_1 and n_1 are, respectively, the concentrations of holes and electrons when the Fermi energy level is at the defect level; Δ_P and Δ_N are the concentrations of holes and electrons, respectively; v_p and v_n are the thermal velocities of holes and electrons, respectively; σ_p and σ_n are the trap-capture cross sections; and N_T is the trap concentration.

Radiative Recombination

In radiative recombination, the electrons and holes recombine, and photons are produced. The carrier-transport process can be expressed as [16]

$$e^- + h^+ \rightarrow hv, \tag{33.6}$$

where e^- represents electrons, h^+ represents holes, and hv represents the photon energy generated.

Auger Recombination

Auger recombination is different from the first two methods, as it is composed of three carriers; that is, when the electron and hole combine to produce energy, the energy is received by a third carrier, which causes the energy absorption carrier to appear in the transition process. According to the carrier difference, Auger recombination involves two types of transition processes: One is the transition of the hole into the valence band, and the other is the transition of the electron into the conduction band; they can be expressed as [17]

$$e^- + h^+ + h^+ \rightarrow h^+ + E_p \tag{33.7}$$

$$e^- + e^- + h^+ \rightarrow e^- + E_p, \tag{33.8}$$

where E_p is the energy required for the third carrier to enter the band edge. Electrons or holes in the valence band that return to the conduction band will participate again in the recombination process.

Electrical Properties of LEDs

An LED is essentially a diode, so its volt–ampere characteristic is determined by the P–N junction of its light-emitting substrate, and its volt–ampere characteristic curve is similar to that of the P–N junction. Figure 33.4 shows the LED operating curve. As shown in the figure, the white LED has one-way conductive characteristics. When the reverse voltage exceeds the reverse voltage of the LED V_C, the LED breaks down, and when the reverse current instantaneously increases, the LED works in the reverse breakdown zone.

When the forward voltage of the LED reaches the operating threshold, V_A, the LED begins to emit light. When the loading voltage is between V_B and V_D, the current I through the LED is proportional to the voltage V at both ends of the LED, and the LED operates in a linear region, as shown by the workspace.

For a P–N junction without a bias voltage, the holes in the P-type doping region and the electrons in the N-type doping region diffuse and form a depletion region, which leads to the generation of a built-in electric field. The potential difference V_D of the built-in electric field is [19] for the P–N junction with the free-hole concentration of N_A in the P-type doping region and the free electron concentration of N_D in the N-type doping region. The potential difference V_D of the built-in electric field is

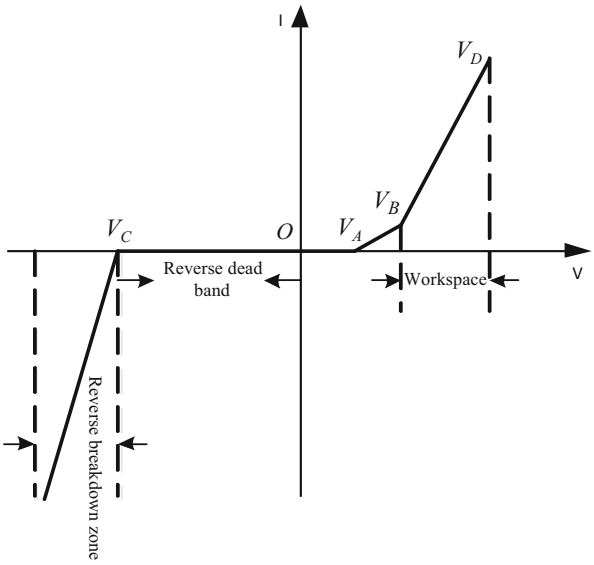


Fig. 33.4 LED operating characteristics [18]

$$V_D = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2}, \tag{33.9}$$

where T is the temperature, q is the charge, k is the Boltzmann constant, and n_i is the intrinsic carrier concentration.

Owing to the presence of a built-in electric field, the movement of the free carriers is limited by the need to overcome the barrier to reach a neutral region with opposite electrical properties. The width of the depletion layer is [20]

$$W_D = \sqrt{\frac{2\varepsilon}{q}(V_D - V)\left(\frac{1}{N_A} + \frac{1}{N_D}\right)}, \tag{33.10}$$

where $\varepsilon = \varepsilon_0 \varepsilon_r$ is the permittivity of the semiconductor materials, ε_0 is the permittivity of vacuum, ε_r is the relative permittivity of the dielectric relative to vacuum, and V is the external voltage applied to the P–N junction. It can be seen from Eq. (33.10) that the width of the depletion layer narrows when V_D approaches V.

A description of the I–V characteristics of the diode was first proposed by Shockley. For a diode with hole diffusion coefficient D_n, electron diffusion coefficient D_p, electron minority lifetime τ_n, and hole minority lifetime τ_p, when the cross-sectional area is S, the I–V characteristics of the diode satisfy the Shockley equation [21]:

$$I = qs\left(\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}}\frac{n_i^2}{N_A}\right)\left(e^{qV/kT} - 1\right). \tag{33.11}$$

At the reverse bias, the diode's I–V characteristics are satisfied, considering the diode's current saturation phenomenon:

$$I = I_{so}\left(e^{qV/kT} - 1\right), \tag{33.12}$$

where the reverse current $I_{so} = eA\left(\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}}\frac{n_i^2}{N_A}\right)$.

For a normal forward bias, $V >> kT/q$, and Eq. (33.11) can be simplified as follows:

$$I = Aq\left(\sqrt{\frac{D_p}{\tau_p}}N_A + \sqrt{\frac{D_n}{\tau_n}}N_D\right)e^{q(V_D - V)/kT}. \tag{33.13}$$

When the applied external voltage is higher than the built-in voltage, the current of the P–N junction increases sharply. The Shockley equation describes the theoretical relationship between the P–N junction voltage and current. In an experiment, considering the presence of ohmic contacts, junction defects, or losses, new formulas are required to describe the I–V properties of actual LEDs:

$$I = I_S e^{eV/(n_{\text{ideal}}kT)}. \tag{33.14}$$

For an ideal diode, $n_{\text{ideal}} = 1.0$. For an actual diode, n_{ideal} is between 1.1 and 1.5. For III–V arsenide and phosphide, n_{ideal} can reach 2.0, and for an InGaN/GaN diode, n_{ideal} can reach 7.0.

In the above model, the parasitic resistance of the device is not considered. If the parallel resistor R_p and series resistor R_S are added, the modified I–V characteristic accords with

$$I - \frac{(V - IR_s)}{R_p} = I_s e^{e(V-IR_s)/(n_{\text{ideal}}kT)}. \tag{33.15}$$

Optical Properties of LEDs

In indoor lighting, the luminous flux Φ (lm) of an LED is expressed as

$$\Phi = kP_{st}, \tag{33.16}$$

where k is the optical efficiency and P_{st} is the radiation power. The range of visible light is from 380 nm to 780 nm, so the above formula can be converted to [22]

$$\Phi = K_m \int_{380}^{780} p(\lambda)V(\lambda)d\lambda, \tag{33.17}$$

where K_m is the maximum visibility, $P(\lambda)$ is the radiation energy spectrum, and $V(\lambda)$ is the luminescence curve.

Researchers have focused on the illumination value, expressed by E (lux), that is, the luminous flux per unit area. Under ideal conditions, the illuminance of a horizontal plane is [23]

$$E = \frac{d\phi}{dA} = \frac{I(\phi)}{D^2}, \tag{33.18}$$

where A represents the irradiated area, D represents the irradiated distance, and $I(\phi)$ represents the light intensity (cd), which is related to the radiation angle ϕ:

$$I(\phi) = kR(\phi), \tag{33.19}$$

where $R(\phi)$ is the radiation intensity model, which satisfies the Lambert model:

$$R(\phi) = aP_t \cos^m(\phi), \tag{33.20}$$

where $a = 2\pi/m + 1$, P_t is the average radiation power, and m is the radiation order, determined by the half-power angle $\phi_{1/2}$, which can be expressed as

$$m = -\frac{\ln 2}{\ln\left(\cos\left(\phi_{1/2}\right)\right)}. \tag{33.21}$$

Combining the equations yields [24]

$$I(\phi) = \frac{m+1}{2\pi}\Phi\cos^{m}(\phi). \tag{33.22}$$

when $\phi = 0$ is introduced into Eq. (33.22), $I(0) = \Phi(m + 1)/2\pi$, where $I(0)$ denotes the central light intensity. Therefore, the intensity of the light source LED can be expressed as

$$I(\phi) = I(0)\cos^{m}(\phi). \tag{33.23}$$

LEDs with different emitting colors typically exhibit different optical properties. To produce visible light, the wavelength should be between 380 and 780 nm, and the peak luminescence wavelength λ depends on the bandgap width of the semiconductor E_g (eV), which is $\lambda \approx 1240/E_g$ [25].

Different colors have different energy barriers; that is, the voltage drop also depends on the color of the light. The on-voltage drop of the red LED is approximately 2 V and that of the blue LED is approximately 3 V. Table 33.1 lists the semiconductor materials corresponding to the light sources of the main colors.

Table 33.1 Light colors corresponding to different LED materials [26]

Color	Lambda wavelength (nm)	Forward voltage (V)	Semiconductor material	Symbol
Red	760 to 610	1.63–2.03	Gallium arsenide phosphide	GaAsP
			Aluminum gallium indium phosphide	AlGaInP
Orange	610 to 590	2.03–2.10	Gallium arsenide phosphide	GaAsP
			Gallium phosphide (doped)	GaP
Yellow	590 to 570	2.10–2.18	Gallium arsenide phosphide	GaAsP
			Indium gallium nitride	InGaN
Green	570 to 500	2.18–4.0	Gallium phosphide	GaP
			Aluminum gallium indium phosphide	AlGaInP
Blue	500 to 450	2.48–3.7	Gallium nitride	GaN
			Silicon carbide	SiC
Purple	450 to 380	2.76–4.0	Indium gallium nitride	InGaN
			Aluminum nitride	AlN

LED Research and Status

LED colors can be divided into red, yellow, green, blue, etc. Light-emitting surfaces can be divided into round, square, and rectangular lamps, surface light-emitting tubes, side tubes, micro-tubes for surface installations, etc. Circular lamps are divided into 2 mm, 4.4 mm, 5 mm, 8 mm, 10 mm, and 20 mm according to their diameter.

Research and Development of Red LEDs

The first practical LED was a red LED [27] developed in 1962. It was made from a composite semiconductor alloy, gallium arsenide phosphide (GaAsP), with the structure shown in Fig. 33.5. The LED's light efficiency was only a tiny 0.1 lm/W, but it set the stage for the future commercial expansion of LEDs. From 1968, when the red LED was made commercially available by Hewlett–Packard (HP), commercial LEDs were used only as indicators until the mid-1990s.

In 1969, gallium nitride (GaN) films were first grown on sapphire substrates using hydride vapor-phase epitaxy (HVPE) by Nakamura. Unfortunately, the crystal quality of these films was very poor [28]. To improve the quality of GaN films, Yoshida et al. used molecular beam epitaxy (MBE) [29] to grow an aluminum nitride (AlN) buffer layer in front of the GaN films.

Amano et al. first used metal–organic chemical vapor deposition (MOCVD) to grow a low-temperature AlN buffer layer and then grew GaN using thin organometallic chemistry [30]. In addition to the AlN buffer layer, Nakamura et al. used a low-temperature (LT) GaN buffer layer to grow GaN films, which produced GaN films with relatively high crystalline quality [31].

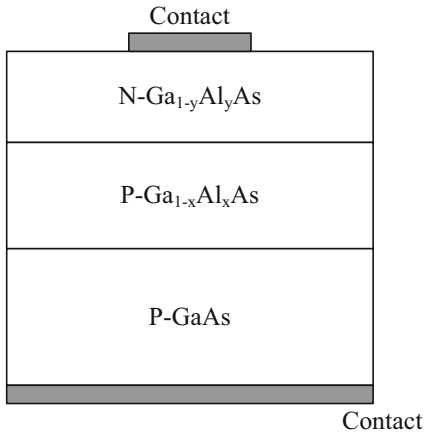


Fig. 33.5 Structure of the first-generation red LED [27]

In 1998, Lumileds Lighting introduced the first commercial high-power red LED [32]. At an input power of 1 W, the working power of the device was 20 times that of the traditional LED lamp, and the efficiency was improved by 50%. More importantly, the thermal resistance of the package was significantly reduced, from 300 K/W to less than the 15 K/W required by Luxeon, and the light efficiency reached 55 lm/W [32].

In 2000, red LED light efficiency reached 100 lm/W, which was sufficient light efficiency to move it gradually from indicator lights to common traffic lights, monitors, etc. Since the end of 2000, Stockholm, Sweden, has completely replaced traffic lights with LED lights, saving 85% of its electricity use, much more than expected [33]. The average power consumption per traffic light decreased from 70 W to 6 W. Simultaneously, Singapore and most parts of China also began to use LED traffic lights [33].

In 2012, Ohkawa et al. proposed a novel high-temperature growth technique to grow high-quality and high In-content indium gallium nitride (InGaN) layers. A large LED with a peak wavelength of 740 nm and electroluminescence (EL) emission was demonstrated.

In 2014, for the first time, Toshiba demonstrated an InGaN-based red LED with a wavelength of 629 nm, an optical output power exceeding 1 mW at 20 mA, and a structure shown in Fig. 33.6. The external quantum efficiency (EQE) was as high as 2.9% [34]. They used an AlGaN intermediate layer in the active region to improve the crystallization quality of InGaN quantum wells and increase the overlap of electron and hole wavefunctions. However, this structure had some disadvantages, such as a high forward voltage that negatively affected the wall-plug efficiency (WPE).

In 2017, to study their optical and electrical properties, Hwang et al. prepared ultrasmall InGaN-based LEDs with an area of 10^{-4} mm^2–0.01 mm^2 [35]. The out-of-peak quantum efficiencies of the smallest and largest LEDs were 40.2% and 48.6%, respectively. The experimental results showed that ultrafine-grained LEDs could be fabricated with a minimal loss of optical or electrical properties.

In 2020, an attractive approach taken by Bi et al. was to prepare high crystal quality InGaN templates using nanostructures, such as nanopyramids and wafers; the structure is shown in Fig. 33.7. These materials have great potential for fabricating

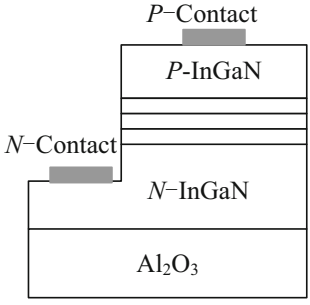


Fig. 33.6 InGaN-based red LED structure [34]

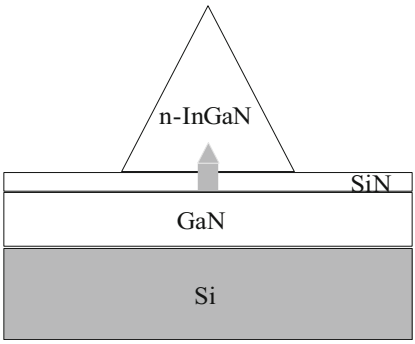


Fig. 33.7 InGaN pyramid growth structure [36]

efficient red LEDs [36]. This structure takes advantage of the strain relaxation, low dislocation density, and reduction of the piezoelectric field. In other words, InGaN quantum wells on these nanostructures can improve the internal quantum efficiency (IQE).

As the luminous efficiency of red LEDs increases annually, they are gradually being transferred to traffic lights, monitors, biomedicine, etc. The external quantum efficiencies of LEDs with different MOCVD sizes are compared in Table 33.2.

Research and Development of Yellow LEDs

The development of the power efficiency of LEDs with various colors has been very uneven, and the efficiency of green and yellow light, which is the most sensitive to light, is much lower than that of blue, purple, and red light in the long term. Studies of yellow LEDs date back to 1965, when Thomas and Hopfield of the American Bell Labs first observed yellow photoluminescence in nitrogen-doped gallium phosphide [53].

The first nitrogen-doped gallium phosphide arsenide P–N yellow LED was fabricated on a GaAs substrate using vapor-phase epitaxy (VPE) technology by Craford of Monsanto Company; the optical efficiency was approximately 0.01% at a current density of 20 A/cm^2 [54]. In 1990, an aluminum indium gallium phosphate double-heterojunction yellow LED with a 574 nm wavelength was grown on a GaAs substrate using MOCVD technology by the Hewlett–Packard company in the United States; at a current density of 30 A/cm^2, the light efficiency reached 0.5% [55]. In 1994, HP transferred an epitaxied AlGaInP yellow LED from a light-absorbing GaAs substrate to a light-transmitting GaP substrate with a 571 nm wavelength and an optical efficiency of 1.3% at a current density of 44 A/cm^2 [56, 57].

In 2008, Nakamura et al. used MOCVD technology to grow InGaN single quantum well yellow LEDs on a semi-polar GaN substrate with a 562 nm wavelength and an external quantum efficiency of 13.4% at a current density of 7 A/cm^2 [58]; however, the voltage was higher and the light efficiency was only 5.5%. In 2013, Toshiba (Japan) used MOCVD technology to grow InGaN yellow LEDs on a sapphire substrate at a 559 nm wavelength, with an external quantum efficiency of

Table 33.2 External quantum efficiencies of LEDs with different MOCVD sizes

Year	Organization	Substrate	Wavelength (nm)	Chip area (μm × μm)	EQE (%)	Reference
1998	Nichia Corp.	Sapphire	594	350 × 350	3.35	[37]
1999	Nichia Corp.	Sapphire	675	350 × 350	0.82	[38]
2002	Nichia Corp.	Sapphire	631	350 × 350	1.1	[39]
2012	Tokyo Univ. of Science	Sapphire	740	500 × 500	2.8	[40]
2014	Toshiba Corp.	PSS	629	460 × 460	2.9	[41]
2015	Chinese Academy of Sciences	Sapphire	585	500 × 1000	1.4	[42]
2016	Tokyo Univ. of Science	Sapphire	600	370 × 370	0.8	[43]
2016	Tokyo Univ. of Science	Sapphire	631	370 × 370	0.4	[44]
2017	Chonbuk National Univ.	PSS	597	1000 × 1000	1.2	[45]
2018	Osaka Univ.	Sapphire	621	–	3.3	[46]
2019	Nanchang Univ.	Si (111)	585	1000 × 1000	12.6	[47]
2020	Univ. of Grenoble Alpes	InGaNOS	612	300 × 300	0.01	[48]
2020	Univ. of Grenoble Alpes	InGaNOS	616	50 × 50	0.09	[48]
2020	KAUST	PSS	633	400 × 400	1.6	[49]
2020	KAUST	β-Ga_2O_3	665	650 × 250	0.19	[50]
2021	UCSB	InGaN	632	6 × 6	0.2	[51]
2021	KAUST	PSS	626	47 × 47	0.36	[52]

18.7% at a current density of 9.5 A/cm^2; owing to high voltages, the light efficiency was only 9.6% [59].

In 2015, Jiang et al. used MOCVD technology to grow yellow GaN-based LEDs on a silicon substrate and used V-shaped-pit multi-quantum well structures, as shown in Fig. 33.8. A yellow LED with a 565 nm wavelength driven by 20 A/cm^2 achieved a light efficiency of 30.7%, corresponding to 164 lm/W [60].

In 2021, Zhao et al. used MOCVD technology to grow a yellow LED with an InGaN-based three-layer staggered quantum well on a sapphire substrate, the structure of which is shown in Fig. 33.9 [57]. The wavelength was 570 nm. At 15 mA, the optical output power and external quantum efficiency of the LED were 10.2 mW and 30.8%, respectively [61].

Research and Development of Green LEDs

In 1991, Nakamura et al. implemented a GaN-based high-power blue LED [62]. Prior to that, light-emitting diodes were dominated by long-wavelength LEDs, such

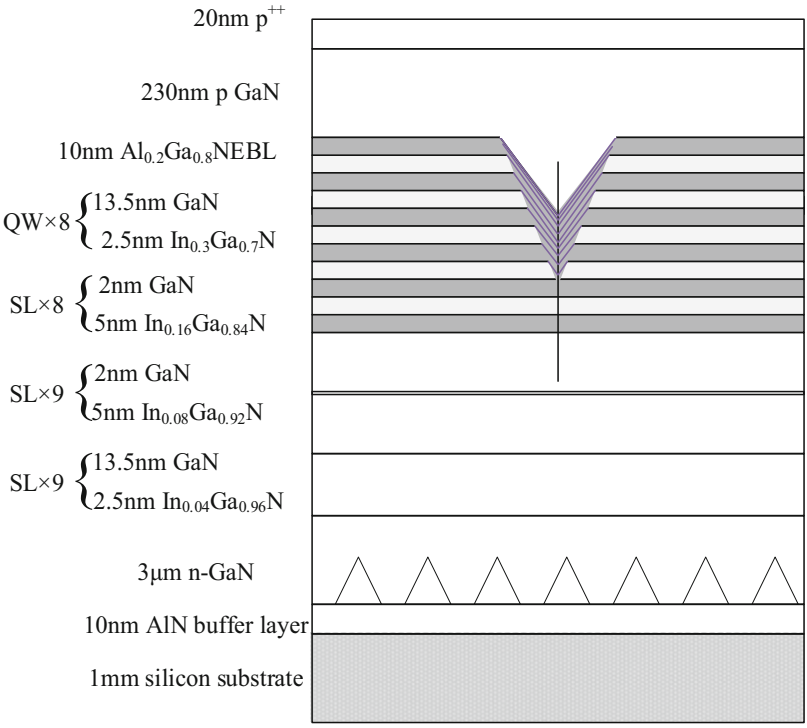


Fig. 33.8 GaN/Si-based yellow LEDs [60]

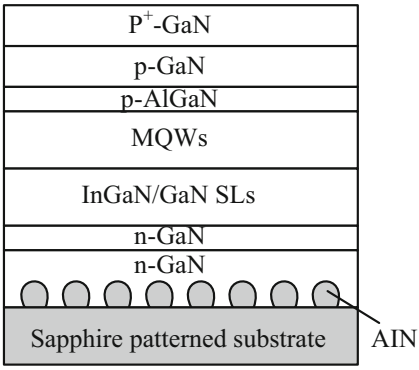


Fig. 33.9 InGaN-based three-layer interleaved quantum well yellow LEDs [61]

as red and orange, in phosphates. Since then, with the rapid development of AlGaN LED systems, short-wavelength ultraviolet LEDs and long-wavelength green and yellow LEDs have been developed. Both the blue LED of the AlGaInN system and the red LED of the AlGaNP system can achieve a high luminous efficiency.

However, LEDs with nitride and phosphide systems cannot achieve high luminous efficiencies in the green and yellow light bands.

In 1998, Lester et al. demonstrated an InGaN-based green LED, with the highest external quantum efficiency (EQE) at the time. The 12-cycle extended InGaN/GaN LED on a sapphire substrate had a multiple quantum well (MQW) structure as the active region, with only 4.5% of its EQE [63]. In 2013, Saito et al. fabricated GaN-based green-yellow LEDs with high efficiency on a C-plane sapphire substrate, using energy-band engineering to design the structure of the device and optimize the growth process; at an operating current of 20 mA, the light output power (LOP) and EQE reached 11.0 mW and 33.7%, respectively, with an almost fivefold increase in efficiency compared with that in 1998 [64].

In 2019, Pli et al. experimented with high-efficiency green LEDs by inserting AlGaN layers in the MQW region, as well as ultrasmall packaging technologies, to improve the EQE and LOP of LEDs [65]. At a current density of 20 A/cm^2, the EQE of the 523-nm nGaN-based green LED was 38%.

Since the twenty-first century, GaN-based green LED research has made significant progress and the luminous efficiency and power have greatly improved [66]. However, compared with blue and red LEDs, green LEDs have obvious disadvantages. In addition to the low quantum efficiency, they have low efficiency under high currents. As the operating current increases, the luminous power of the LED increases, accompanied by a decrease in luminous efficiency; this is called the droop effect [67].

Perovskite light-emitting diodes (PeLEDs) are light-emitting diodes with high color purity, tunable wavelengths, and low production costs. They have been studied extensively, and their EQEs have continued to rise, as shown in Fig. 33.10.

The PeLED device has a sandwich structure, in which the light-emitting layer is sandwiched between the electron injection layer (EIL) and the hole-injection layer (HIL) [69]. Depending on the order of the implanted layers, the specific device structures can be further divided into two types: positive and inverted.

In the positive structure (Figure 33.11a), the transparent conductive glass substrate acts as the positive electrode of the device, with the HIL, perovskite layer, EIL, and metal electrode (negative electrode) placed on top. In the inverted structure (Figure 33.11b), the transparent conductive glass substrate acts as the negative electrode of the device, with the EIL, perovskite layer, HIL, and metal electrode (positive electrode) placed on top [69].

The first green PeLED (GPeLED) device was developed in 2014 by Friend et al. at the University of Cambridge. They demonstrated a prototype device with an EQE of 0.1% and luminance of 364 cd/m^2; the structure is shown in Fig. 33.12. Shortly thereafter, the Greenham group produced a GPeLED with a 1% EQE by adding an additive to the precursor solution to inhibit crystal growth [70], early in 2015. Huang and Wang et al. lowered the electron injection barrier by adding a polyethylenimine buffer layer to increase the brightness of the GPeLED to over 20,000 candela per square meter (cd/m^2) [71].

At the end of 2015, Lee et al. [72] elevated the EQE of GPeLEDs to an unprecedented 8.53% by nanocrystalline pinning and changing the composition

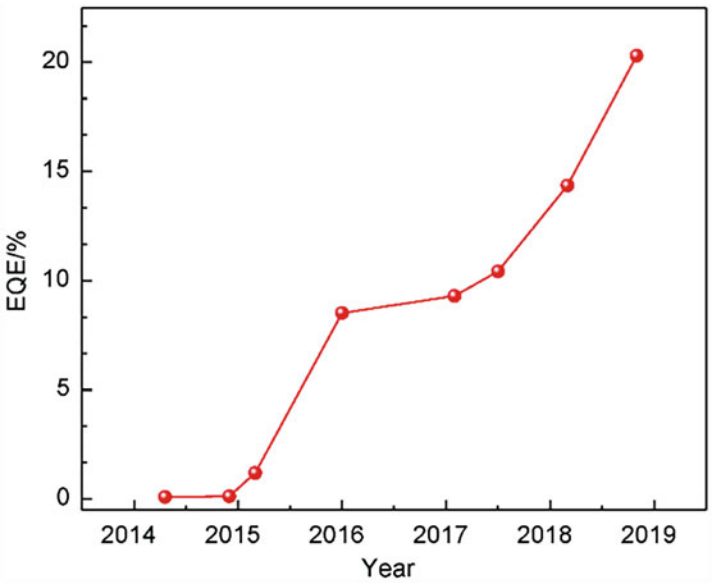


Fig. 33.10 Green PeLEDs' EQE growth trend [68]

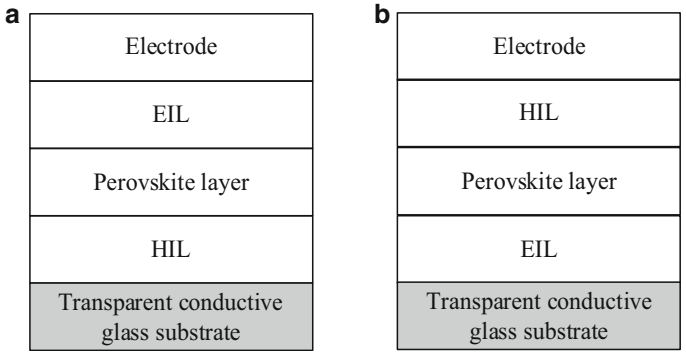


Fig. 33.11 Typical perovskite LED structures. (**a**) Positive structure (**b**) Inverted structure

ratio of the precursor solution. As the EQE continued to rise, the brightness of the GPeLED increased to a stunning 591,197 cd/m^2 [73]. In early 2017, Rand et al. produced GPeLEDs with an EQE of up to 9.3% [74]. In the same year, You et al. increased the EQE of GPeLEDs by 10% and then advanced it to nearly 15% [75]. Recently, the research results of Wei et al. greatly improved the EQE of GPeLEDs, breaking the 20% milestone, and laying a solid foundation for the commercialization of GPeLEDs [76].

Because green is one of the three primary colors (red, green, and blue), it is very important to produce high-efficiency green LEDs, whether to achieve high-

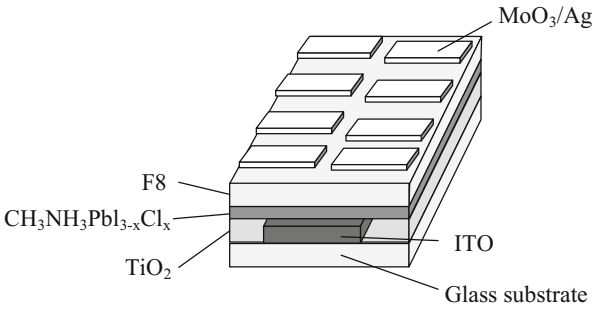


Fig. 33.12 GPeLED structure diagram

efficiency multi-basic white lighting or full-color displays with high resolution and fast reaction speeds.

Research and Development of Blue LEDs

Blue is a basic color in nature and one of the three primary colors of the plane panchromatic display. At first, SiC and ZnSe were the important materials for blue LEDs; GaN was neglected because of the lack of a single-crystal substrate, high dislocation density, and P-type doping. However, with continuing research and improving technology, GaN-based blue LEDs have become mainstream research. The parameters of GaN are listed in Table 33.3 [77].

The first blue GaN-based LEDs were introduced in 1971 [78]. Because GaN P-type doping was not possible at that time, a metal–insulator–semiconductor (MIS) structure was adopted, the basic structure of which is shown in Fig. 33.13. The LED luminescence efficiency of the MIS structure was only 0.03%–0.1%. The peak wavelength was approximately 485 nm, the full width at half maximum (FWHM) was 70 nm, the typical operating voltage (when the input current was 20 mA) was 7.5 V, with a light output of 2 mcd at 10 mA; however, it had a longer lifetime [78].

In 1983, Yoshida et al. deposited a layer of AlN as a buffer layer on a sapphire substrate, which greatly improved the surface structure and crystal quality of GaN P-type LEDs. The doped GaN was obtained by Mg doping and low-energy electron beam irradiation (LEEBI). These two breakthroughs laid the foundation for the generation of GaN-based LEDs [79].

In 1991, Nakamura et al. developed the Shuangliu District organometallic chemistry vapor deposition technique, which uses vertically injected $N_2 + H_2$ gas streams. High-quality GaN films were successfully grown on a low-temperature GaN buffer layer by suppressing the thermal-convection effect of the reactive gas flow (TMG + $NH_3 + H_2$) on a sapphire substrate at temperatures up to 1000 °C [80]. In addition, they developed a technique for obtaining high-quality P-type GaN materials by annealing in clean N_2 gas.

Table 33.3 GaN material characteristic parameters [77]

Characteristics of wurtzite GaN	Bandgap width ($T = 300$ k)	Lattice constant/nm	Coefficient of thermal expansion ($T = 300$ k)	Thermal conductivity/ $W \cdot cm^{-1} \cdot K^{-1}$	Refractive index	Electronic effective mass
	$E_g = 3.39$ eV	$a = 0.3189$	5.59×10^{-6} K	$K = 1.3$	n (1 eV) $= 2.33$	$m_c = 0.20$ $m_0 = 0.02$
		$c = 0.5185$	3.17×10^{-6} K		n (3.38 eV) $= 2.67$	
Characteristics of sphalerite GaN	Bandgap width ($T = 300$ k)	Lattice constant/nm		Donor–recipient peak energy level	Free electron-acceptor value energy level	
	$E_g =$ 3.29–3.35 eV	$A = 0.452$–0.455 nm		3.196 eV	3.262 eV	

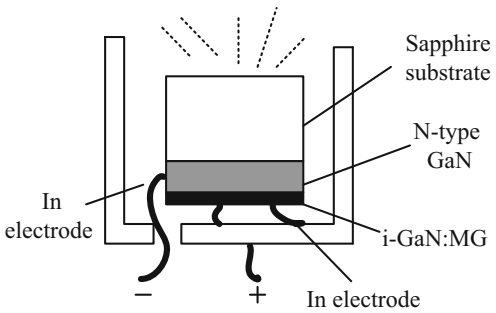


Fig. 33.13 GaN-based MIS epitaxial layer and device structure [78]

In 1993, Nakamura et al. began to commercialize their research and development of GaN-based blue light LEDs [81]. At the end of 1994, a high-brightness GaN-based blue LED device was developed [82]. Figure 33.14 shows the structure of the LED, which won the Nobel Physics Prize in 2014. The device has a double heterostructure, and Zn-doped InGaN is used as the active layer for emitting blue light.

In 2000, Li et al. of Kansas State University fabricated GaN quantum well micro-LED arrays with a single 12-μm diameter and a 10 × 10 array via photolithography and plasma dry etching. This shows that LEDs were moving toward smaller sizes. The results showed that the current density of micro-LEDs was five times higher than that of large LEDs at the same voltage [83].

In 2004, Choi et al. used a photoresist as a mask to etch GaN using inductively coupled plasma (ICP) with different inclined angles by adjusting the process parameters in the ICP dry etching process An InGaN-based micro-LED array device with passive driving capability was fabricated, which had a resolution of 128 × 96 [84].

In 2008, using Ni/Au/Ti/Au as a p-plane mirror and ohmic contact layer, Gong et al. of the University of Schottfried, UK, fabricated GaN epitaxial films with a 20-μm pixel size and micro-LED arrays in blue (470 nm) and purple (370 nm) with a

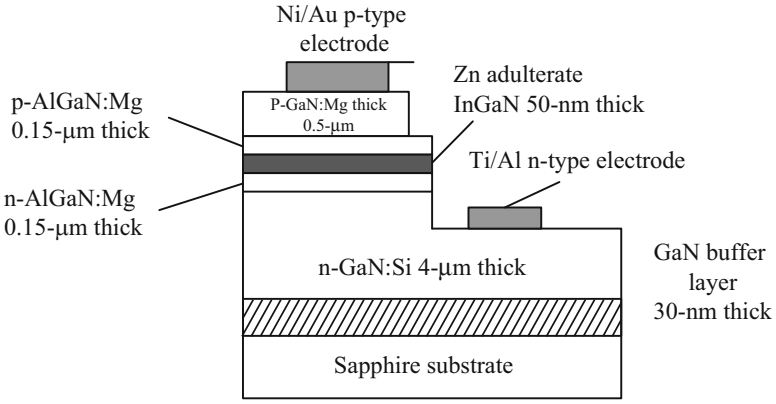


Fig. 33.14 Double-heterojunction GaN-based blue LED structure [82]

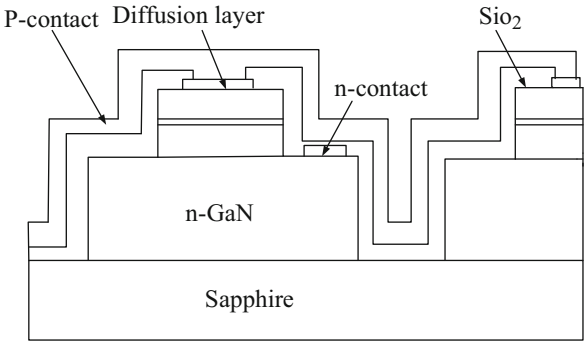


Fig. 33.15 Micro-LED with an inverted structure [85]

50 μm pixel center distance. The array device had 64 × 64 pixels, and the structure is shown in Fig. 33.15 [85].

In January 2014, Tian et al., from the University of Schleicher, UK, used metalorganic vapor-phase epitaxy (MOVPE) to grow GaN epitaxy on a six-inch Si (111) substrate. A GaN blue micro-LED array with passive driving capability was fabricated by layering the positive and negative electrodes. The number of pixels of the array was 10 × 10, the pixel size of each micro-LED was 45 μm, and the central wavelength was approximately 470 nm; a schematic of the array device structure is shown in Fig. 33.16 [86].

In March 2018, Lei et al. of PKU Blue Bird Display, Ltd. proposed wafer-level hybrid integration technology and fabricated three types of red, green, and blue micro-LED array display devices with a single-color active drive [87]. This wafer-level hybrid integration technology is different from flip-chip welding, which requires bonding the entire wafer to the CMOS drive circuit and then removing the substrate of the epitaxial wafer. A microdisplay device is fabricated on a CMOS driver using a standard semiconductor process. The structure is shown in Fig. 33.17 [87].

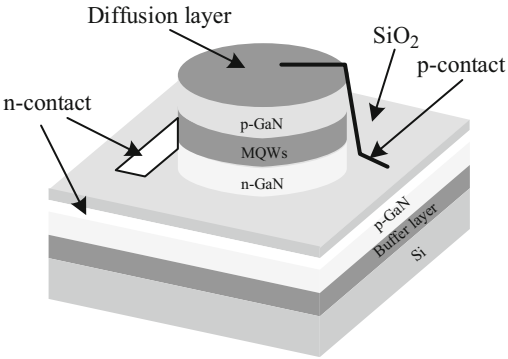


Fig. 33.16 Micro-LED structure on silicon substrate [86]

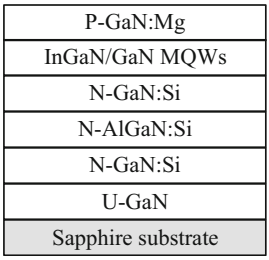


Fig. 33.17 Sapphire substrate micro-LED structure [82]

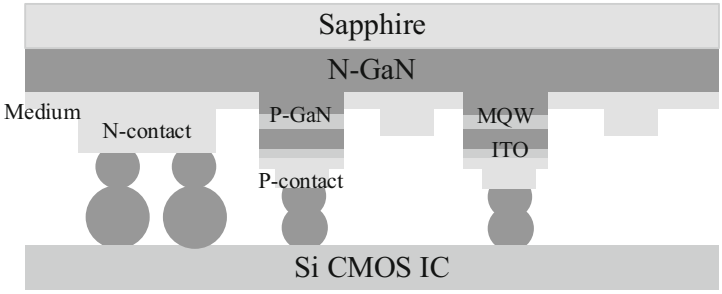


Fig. 33.18 Flip-chip micro-LED structure [88]

In April 2019, Chen et al. of National Tsing Hua University fabricated an inverted GaN-based blue light micro-LED M-array display device driven by CMOS via In–In bonding. The number of device arrays was 960 × 540, the pixel size was 8 μm, and the pixel center interval was only 12.8 μm. The test results showed that the leakage current of the device was only 0.1 pA at −5 V, and the open voltage was 2.73 V. The overall pixel yield of the display device reached 98.67%. A schematic is shown in Fig. 33.18 [88].

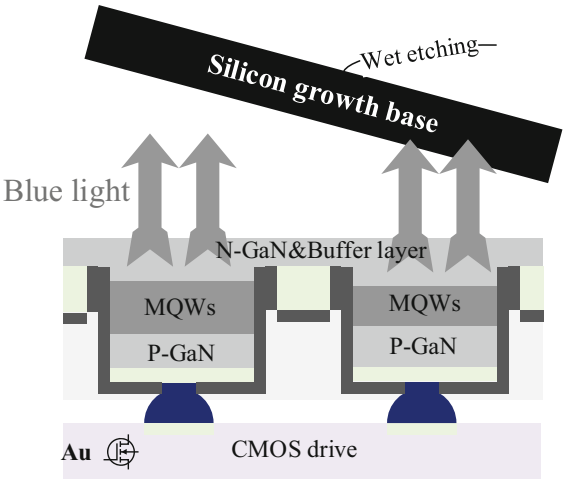


Fig. 33.19 Micro-LED structure on a flip-chip silicon substrate [89]

In March 2021, Qi et al. of the Hong Kong University of Science and Technology etched and fabricated an inverted micro-LED array on a blue GaN epitaxy on a silicon substrate and fabricated an Sn column on the anode back into an Sn sphere. The micro-LED array was combined with a CMOS driving backplane by Au–Sn bonding. The silicon substrate was then removed by wet etching, and the micro-LED microdisplay device was successfully fabricated. The microdisplay device had a resolution of 400 × 240 pixels, and the overall pixel yield was 95%. The single pixel size was 20 × 20 μm^2, the center distance of the pixel was 30 μm, and the number of pixels on the diagonal line per inch was 848. The structural principle is illustrated in Fig. 33.19 [89].

Research and Development of White LEDs

High-efficiency purple, blue, and green (GaInN LED) devices were developed in the early 1990s [90]. Since the mid-1990s, monochrome high-efficiency devices have been able to cover the entire visible spectrum with the necessary conditions to achieve white LEDs. The white LED principle and the monochrome LED principle are different; for white LEDs, there are several feasible methods.

The first method involves mixing red, green, and blue LED chips, in which the three LED chips emit the three primary light colors mixed together to produce white light [91]. This method is shown in Fig. 33.23. As can be seen from the diagram, the LED requires four lead electrodes to ensure that each of the three LED chips can be injected with the appropriate current. With the advent of efficient blue and green LEDs, this approach has become more feasible [92].

Nakamura et al. [93] showed that a combination of high-power and high-brightness blue InGaN SQW LEDs, green InGaN SQW LEDs, and red GaAlAs

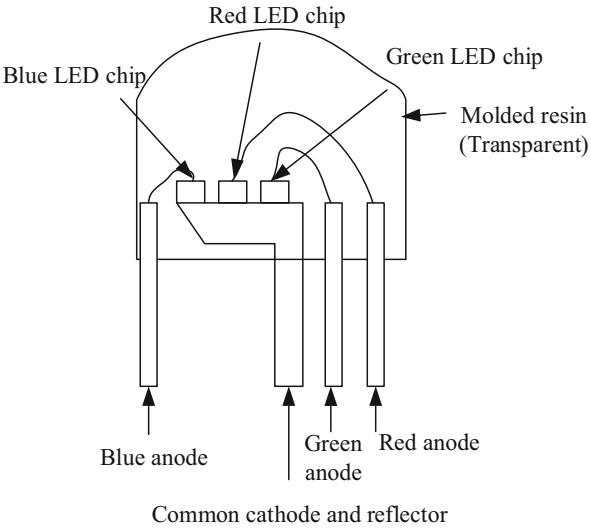


Fig. 33.20 White LED made of red, green, and blue LED chips

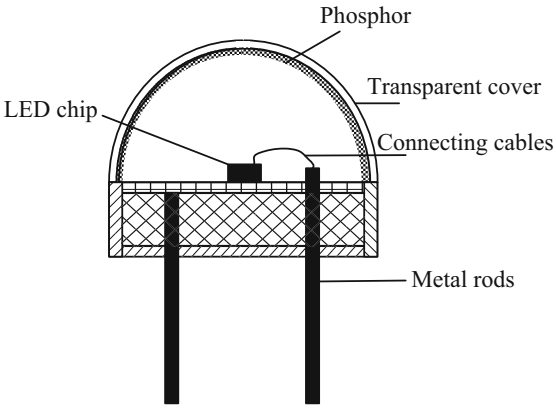


Fig. 33.21 UV-excited white LED

LEDs could be used for a variety of applications, such as LED full-color displays, and used to replace light bulbs or fluorescent LED white lamps (Figs. 33.20 and 33.21).

The second method is a white LED excited by ultraviolet (UV) light, which is used to excite red, green, and blue composite fluorescent materials by near-UV rays [72]. Because it is similar to conventional fluorescent lamps [95], this approach was proposed many times before the advent of GaN-based blue LEDs [96]. The LED structure is shown in Fig. 33.24. An LED chip and phosphor coating inside the transparent housing form a point-source white LED.

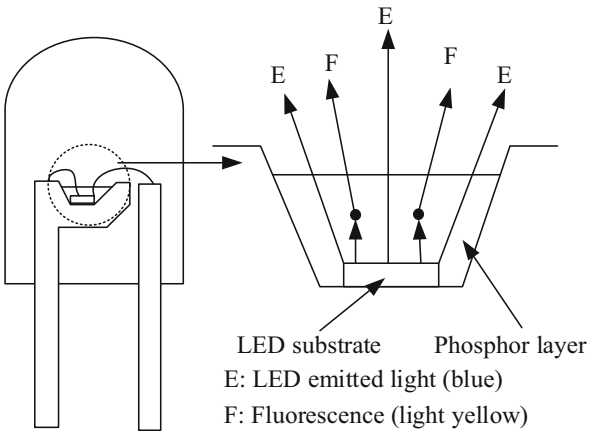


Fig. 33.22 White LED with a blue light and YAG phosphor

The third and most successful method is the combination of a blue LED chip and phosphor [97]. The principle is that the blue light from the LED chip is absorbed by the phosphor, whereas the other part of the blue light is transmitted through the phosphor. Thus, blue light (from LED chips), green light (from phosphors), and red light (from phosphors) together form white light [98]. The structure is shown in Fig. 33.26.

After realizing that organic phosphors could not reliably operate under the harsh operating conditions of blue LEDs, Shimizu et al. identified a very suitable inorganic phosphor, cerium-doped yttrium aluminum garnet (YAG:Ce), which absorbs blue light and emits red and green light [99]. This method was first demonstrated by Shimizu et al. and further improved by Bando et al. Nakamura and Fasol subsequently reviewed it. This approach has been commercially successful because it is efficient, simple, and requires only one LED chip to be powered by a single power supply (Fig. 33.22).

The last and most difficult of the three primary LEDs, blue, has been present since Nakamura developed a bright blue LED in 1994. Subsequently, LEDs gradually moved toward the white lighting stage.

In 1996, Shimizu [100], a Japanese company, published a Japanese patent application for a "Sheet Light Source," an RGB white backlight for full-color displays; it disclosed a flat white light source based on a blue LED and red and green phosphors. In the same year, an article in Nikkei Shimbun, a Japanese newspaper [101], announced a new white light source, based on blue LEDs, that excited a yellow YAG:Ce phosphor. According to reports, the white LED has high efficiency (5 lm/W), low cost, and a life expectancy of 50,000 h.

In 1997, Osram–Fraunhofer et al. [102] reported a combination of blue light GaN chips with luminescent down-conversion inorganic YAG:Ce phosphors and organic fluorescent dyes. The team demonstrated the use of LED chip–phosphor combinations to emit white light and blend colors, such as cyan and magenta.

Traditional LEDs are commonly used in lights, signs, and displays. Before 2000, the general lighting of homes, offices, and streets was not the primary use of LEDs. In 2000, LEDs with high efficiency, low energy consumption, durability, and other characteristics gradually became the future direction of lighting development. Moreover, LED power has rapidly developed, but also meets basic lighting requirements.

In 2004, Narukawa [103] reported that Nichia had broken the 100-lm/W barrier and its luminous efficiency was estimated to exceed 100 lm/W in the low current region (less than 1 mA), indicating that InGaN-based white LEDs had great potential to replace fluorescent lamps (75 lm/W) in the near future.

In 2006, Narukawa et al. [104] demonstrated efficiencies of up to 174 lm/W at a low current of 2 mA and up to 138 lm/W at a current of 20 mA (current density of 1.98 A/cm^2).

In 2010, Mukai, of Nichia, developed a white LED with very low current injection [105], resulting in a laboratory efficiency of 249 lm/W. In the same year, Cree announced that the correlated color temperature of white LEDs was approximately 4600 K at an injection current of 350 mA [106] and the laboratory result was 208 lm/W.

In 2014, Cree released laboratory results for a white LED lamp (excluding power) with a color temperature of 5150 K and an efficiency of 303 lm/W at an injection current of 350 mA [107].

LEDs in the general lighting and functional lighting fields continued to expand and the market size maintained rapid growth. From 2015 to 2020, the global LED market grew rapidly, and LEDs have gradually replaced incandescent lamps, fluorescent lamps, and other lighting sources, with the penetration rate continuing to rapidly increase [108].

By the end of 2020, the global LED lighting penetration rate had exceeded 60%. Using combined data from IHS Markit and LEDinside, the global market for LED devices was approximately $33.17 billion in 2019, up nearly 34.8% from $15.71 billion in 2015 (an average annual growth rate of 7.7%). LED devices for general lighting increased by 38.8% (8.1% average annual growth) from $4.62 billion to $6.41 billion, while those for functional lighting increased from $8.63 billion to $10.21 billion, an increase of 18.3% (average annual growth rate of 4.3%). The market share of LED devices for innovative applied lighting expanded from 3.7% to 5.2% [109].

The improvement in the efficiency of visible light LEDs is truly impressive. Assuming constant progress from 1960 to 2000, the LED efficiency has doubled every 4 years. This phenomenon is consistent with Moore's law [109].

Research and Development of UV LEDs

Ultraviolet light-emitting diodes (UV LEDs) are energy-efficient, environmentally friendly, and long-lived and have significant applications in fields such as sterilization and medical and biochemical testing. In recent years, increasing attention has

been paid worldwide to ultraviolet semiconductor optoelectronic materials and devices; in fact, they have become a research hotspot [110].

Research into AlGaN ultraviolet LEDs began in the 1990s. In 1992, with the realization of P-type doping of AlGaN materials, two researchers, Akasaki and Amano, from Nagoya University, Japan, used the double-heterojunction structure to prepare GaN/AlGaN ultraviolet LEDs for the first time, with a luminescence wavelength of 372 nm [111].

In 1998, in the American Sandia National Laboratories, Han et al. grew a GaN 0.2 GaN MQW epitaxial structure on a sapphire substrate using MOCVD, and successfully developed a UV LED with an MQW structure. The light-emitting wavelength was 353.6 nm, with an output power of 13 W, a start-up voltage of 4 V, and an EQE of less than 1% driven by a current of 20 mA. This was the first-ever development of the shortest-wavelength nitride material LED [112].

To reduce the piezoelectric effect of quantum well materials and improve the EQE of LEDs, in 1999, Nishida et al., [113] in NTT Fundamental Research Laboratory, Japan, used an Al0.06 GaN/Al0.12 GaN MQW structure to suppress the piezoelectric field and created an ultraviolet LED with a 346 nm wavelength.

In 2000, Kinoshita et al. of RIKEN, Japan, developed an LED structure epitaxy growth on a 6H-SiC (0001) substrate using a metalorganic vapor-phase epitaxy (MOVPE) technique. The hole-injection efficiency of an Al0.03 GaN/Al0.25 GaN MQW was increased by using an Mg-doped GaN/Al0.25 GaN superlattice, and the electroluminescence wavelength of the LED was 333 nm at a pulse current density of 0.13 kA/cm^2 [114].

In 2001, Nishida et al. proposed that an output power and internal quantum efficiency (IQE) of over 3 mW and 80%, respectively, could be obtained using bulk GaN as the substrate with a peak wavelength of 352 nm and 100 mA current injection, with an external quantum efficiency of 1%. Under a 400 mA high current injection, the maximum power exceeded 10 W, and the operating voltage was less than 6 V [115].

In 2006, Taniyasu et al. reported an AlN homojunction LED with a p-type/intrinsic/N-type (PIN) structure, which shortened the luminescence wavelength to 210 nm with an external quantum efficiency of 0.1% [116].

In 2009, Hirayama et al. of RIKEN, Japan, epitaxied AlGaN MQW LEDs on a sapphire substrate using an ammonia pulse multilayer growth technique. The maximum optical power was 0.15 mW, and the EQE was 0.2% [117].

In 2010, Nichia, Japan, used a high-quality AlN template to prepare deep-ultraviolet LEDs with a wavelength of 281 nm and an EQE of 2.78% at a driving current of 20 mA [118].

In 2013, deep-ultraviolet LEDs were grown on AlN substrates using HVPE technology by the Tokuyama Company of Japan. At 268 nm, the EQE was 2.4% (current 250 mA); at 261 nm, the EQE was 2.4% (current 250 mA), and the EQE was 2.2% (current 20 mA) [119].

In 2014, UV LEDs were fabricated by RIKEN, Japan, using the multi-quantum barrier (MQB) technique, with an EQE of 7% at 279 nm and a current of 25 mA [120]. In the same year, the UV Craftory in Japan produced UV LEDs with a

wavelength of 285 nm and an EQE of 14.3% at a driving current of 2 mA, using a novel packaging process [121].

In 2015, UV LEDs were fabricated by the US Sensor Electronics Technology company using an encapsulation process with excellent heat dissipation; the light-emitting wavelength was 278 nm, and the EQE was 11% at a working current of 10 mA [122].

In 2017, the Nippon Institute of Intelligence and Communications Technology (NICT) used nanoimprint lithography to fabricate nanoscale surface structures for photon extraction, enabling LEDs to emit light at a 265 nm wavelength and an 850 mA current with an EQE of 3.9%. The continuous output power was 150 mW, which was the highest continuous output power when the LED emission wavelength was below 280 nm [123].

In 2018, Zhang et al. designed an AlGaN-based deep-ultraviolet light-emitting diode with almost no efficiency drop, using a specially designed P-type electron barrier in a superlattice with a peak wavelength of 270 nm; the external quantum efficiency was 25% when the injection current density was 40 A/cm^2 [124].

In 2020, Ha et al. achieved a significant improvement in AlGaN deep-UV light-emitting diode efficiency via local surface plasmon resonance coupling mediated by high-density Al nanoparticle arrays. At an operating temperature of 300 K, the internal quantum efficiency was 12.3%, which was 100% more efficient than conventional UV LEDs [125].

In 2020, Yuan et al. fabricated an InGaN/GaN/AlGaN/GaN MQW UV LED with an output wavelength of 395 nm, an external quantum efficiency of 60%, and an output power of 659 mW at an injection current of 350 mA [126].

LED Device Material History

Before the advent of LEDs, owing to the development of semiconductor technology, there was complete theoretical knowledge explaining the phenomenon of electroluminescence, which laid the theoretical foundation for the development of LEDs. Since 1962, when the first practical red LED was developed, there has been intense discussion about the structure to be used to improve the luminous efficiency of LEDs.

First, SiC and ZnSe were important materials for the research and development of optoelectronic devices; GaN was put on hold because of the lack of a suitable single-crystal substrate, high dislocation density, and a complex fabrication process. The first GaN LEDs were not available until 1971; however, because P-type GaN doping was not possible at that time, only the metal–insulator–semiconductor (MIS) structure was used. Owing to the small number of carriers in the structure and the low efficiency of radiative recombination, the luminescence efficiency was very low, only 0.03–0.1 lm, with an optical output of 2 mcd at a current input of 10 mA [127].

In 1983, Yoshida et al. deposited a layer of AlN as a buffer layer on a sapphire substrate and obtained P-type doped GaN materials via Mg doping and low-energy

electron beam irradiation. This breakthrough laid the foundation for the subsequent generation of P–N junction LEDs [79].

In the late 1980s, the first GaN P–N junction LED was produced at Nagoya University by Amano et al., based on the work of Yoshida et al. n-GaN thin films were grown on an AlN buffer layer using the MOCVD method. The thickness of the N-GaN thin films was only 3 μm; however, the carrier concentration was 2×10^{-17} cm^{-3}. The experimental results showed that the performance of an LED with a P–N junction was better than that of an LED with an MIS structure [128].

Since then, the use of GaN-based LEDs has grown rapidly. In 1991, Nakamura et al. of Nichia Company successfully developed Mg-doped GaN blue LEDs, and for the first time, GaN thin films were grown on a GaN buffer layer by double-flow MOCVD. This greatly improved the film quality, making the hole mobility up to 600 cm^2 V^{-1} s^{-1}; the hole concentration of Mg: GaN was 3×10^{-18} cm^{-3}. The peak wavelength of the light emission was 430 nm, FWHM was 55 nm, light output power was 42 μW ($i = 20$ mA), operating voltage was only 4 V, and the external quantum efficiency was approximately 0.18% [129]. The spectrum quality was good with only one peak.

In homojunction LEDs, owing to the low barrier height, the restrictions on the injected carriers are weak, and the minority carriers are distributed over a wide range; these are orders of magnitude of several diffusion lengths, generally a few microns. In this range, the recombination efficiency of electron and hole radiation is low, so the luminescence efficiency is also low. Heterojunction LEDs were developed to increase the barrier height and enhance the recombination of electrons and holes.

In 1992, Nakamura et al. developed the first p-GaN/n-InGaN/N-GaN double-heterojunction blue LED. The peak wavelength of the output light was 440 nm, and the FWHM was 180 nm. The output power was 125 W, the operating voltage was 19 V when the input current was 20 mA, and the external quantum efficiency was 0.22% [131].

To strengthen the recombination of electrons and holes, a quantum well structure was created; it can be divided into single quantum well and multiple quantum well structures (MQWs). A single quantum well structure, such as a sandwich, consists of two broad bandgap materials sandwiched together. The main principle of MQWs is to introduce a multi-heterostructure as the active region of the LED and increase the probability of carrier trapping to improve the efficiency of radiative recombination.

In 2001, Lai et al. fabricated blue LEDs with an MQW structure. The active layer was composed of five layers of 2.5-nm-thick In0.25 Ga0.75 N and 4-nm-thick GaN, the output wavelength was 445 nm, and the FWHM was 28 nm. The input current was 20 mA, the output power was 2.2 mW, the saturation current was 1.4 A, the output power was 2.2 mW, the output power was 53 mW, and the external quantum efficiency was 4.5% [137].

Heat dissipation is a major obstacle to improving the efficiency of LEDs. Hence, engineers have developed flip-chip structures that place low heat dissipation sapphire substrates on top; heat dissipation materials are added to the substrate to enhance the heat dissipation function. In 2003, there were many patent applications for packaging technology for flip-chip LEDs.

Table 33.4 External quantum efficiencies of LEDs with different structures

Year	Author	LED structure	Wavelength	EQE (%)	Reference
1971	Ponce	MIS	485	0.1	[127]
1989	Amano	P–N junction	370	–	[128]
1991	Nakamura	Homogenous junction	430	0.18	[129]
1991	Sugawarah	Double heterojunction	620	1.5	[130]
1992	Huang	Double heterojunction	555	6	[131]
1993	Nakamura	Double heterojunction	440	0.22	[132]
1994	Mukai	Double heterojunction	500	2.4	[133]
1995	Nakamura	Single quantum well	450	9.2	[134]
1999	Wang	Single quantum well	530	–	[135]
2000	MODAKP	Multiple quantum wells	650	4.35	[136]
2001	Lai	Multiple quantum wells	445	4.5	[137]
2001	Tadatomok	Multiple quantum wells	382	24	[138]
2002	Koike	Multiple quantum wells	485	20	[139]
2006	Shchekin	Flip structure	441	38	[140]
2019	Zhous	Flip structure	460	46.8	[141]

In 2006, Shchekin et al. fabricated a thin-film flip-chip InGaN/GaN multi-quantum well LED with an output wavelength of 441 nm. It had a luminance of 191 mW/mm^2 and an external quantum efficiency of 38% under a driving current of 1 A [140].

These typical structures represent the development of diodes, from the MIS structure to the P–N structure to homogenous and heterogenous structures and then to quantum well and flip-chip structures. The output power and external quantum efficiency of the LEDs were improved by changing the structure. Table 33.4 shows the external quantum efficiencies of LEDs with different structures.

Visible Light Full-Duplex Communication

Research and Development of Visible Light Communication

Visible light wireless communication technology was first proposed by Professor Harald Haas, a German physicist. It is a green information technology that can realize high-speed data transmission by utilizing the fast response characteristic of LEDs. It is known in academia as visible light communication (VLC) technology [142]. As a type of access mode that broadens the resource spectrum, saves energy, and can be moved, it introduces a new idea for the development of traditional RF communication technology. Visible light communication uses LEDs as the light source, which enables it to communicate at high speeds [142].

In 1979, Gfeller created an indoor infrared communication system model, which divided indoor channels into line-of-sight (LOS) and non-line-of-sight (NLOS)

channels. Their impulse responses were calculated to reflect indoor channel characteristics [143].

In 1993, Barry proposed an indoor infrared communication channel model based on Lambert radiation by using a recursive loop algorithm to calculate the channel impulse response [144].

In 2000, Tanaka and Haruyama first proposed a VLC system. They presented the feasibility of using a white LED as the emitting light source of an indoor optical communication system and built a VLC system model [145]. In this study, the indoor VLC system was analyzed mathematically, the VLC channel was studied, and the problems of direct line-of-sight and scattering links were explained in detail.

In 2003, Komine et al. studied the channel model, light source characteristics, and noise model of white LED communication systems and obtained the bit error rate (BER) performance of different power ranges and modulation modes [146].

In 2007, Ke and Ding et al. of Xi'an University of Technology began to study the optimal layout of a VLC light source and designed a set of visible light wireless local area networks (LANs) [147].

In 2008, Arikan proposed channel polarization-based polarimetric codes, which are the only ones that can reach the Shannon limit. The coding algorithm for channel polarization performs well in binary discrete memoryless input channels [148].

In 2009, after extensive research, Komine et al. proposed an effective method to improve communication quality, which combined a decision feedback equalizer with least mean squares (LMS); thus, an adaptive equalization process was adopted for the VLC channel [149]. When the transmission rate was low, both the decision feedback and linear equalization had better effects. However, when the transmission rate was greater than (or equal to) 700 Mbit/s, the interference result of the former was better than that of the latter.

In 2010, Ke and Ding et al. created a new model of the visible light channel and proposed a fast simulation algorithm for light tracking [150].

In 2013, Bandara studied a recursive least-squares (RLS)-based VLC system that reduced the training sequence through theoretical analysis and the numerical simulation of an RLS-based adaptive equalization algorithm [151]. In the same year, He et al. developed a channel model and studied the path from communication access points to LED array access points and the path from there to each LED unit. The modeling effect was closer to reality [152].

In 2016, Li et al. used pulse amplitude modulation (PAM) modulated signals for transmission, based on preequalization. Experimental results showed that PAM signals had higher spectral efficiency than non-return-to-zero (NRZ) signals [153].

In 2017, Fang et al. provided polarimetric code-based VLC dimming control of an efficient forward error correction (FEC) coding scheme [154]. In the same year, Kim et al. studied a polarimetric code dimming control system for VLC [155] and, in 2018, analyzed the impact of reducing the intersymbol interference (ISI) caused by the multipath delay of the signal with polarimetric codes. Yao proposed a joint beamforming and DC bias design for multiuser, multi-input, multi-output (MU-MIMO) VLC systems that minimized the impact on lighting requirements [156].

In 2019, Haas achieved a real-time transmission rate of 15.73 Gbit/s in VLC using orthogonal frequency-division multiplexing (OFDM) based on a four-color wavelength-division multiplexing system [157].

In 2020, Wu et al. proposed a deep neural network (DNN) model for channel estimation based on DC-biased optical orthogonal frequency-division multiplexing (DCO-OFDM) [158]. In this scheme, the transmitted signals are used as tags and the received sampled signals are used as channel features for learning. The DNN model had a better BER performance than the traditional least-squares and least-square-error algorithms and eliminated the complexity caused by the minimum mean-squared error algorithm.

In 2020, Wang et al. proposed a VLC transceiver system that used three independent drivers with a feed-forward equalizer (FFE) to drive RGB LEDs to generate PAM-4 signals. The receiver used a cascading continuous-time linear equalizer (CTLE) as a post-equalizer. This VLC system increased the maximum data rate from 28 Mbit/s to 75 Mbit/s [159].

LED VLC has been widely studied and supported by many countries because of its advantages. To realize high-speed and long-distance transmission of visible light, researchers have thoroughly explored and researched advanced modulation technology, digital signal pre-/post-equalization technology, system transceiver devices, etc., and made a series of breakthroughs. Table 33.5 summarizes some important research results in the field of visible light communication at home and abroad.

Overall Full-Duplex Communication System Structure

As shown in Fig. 33.23, the indoor full-duplex communication system with a single visible light source consists of two parts: the active end and the reverse end. The optical communication link from the active end to the reverse end is the downlink, and the reply communication link from the reverse end to the active end is the uplink.

In this system, the active end is typically fixed to the ceiling of the room to provide lighting while sending information. The reverse terminal is located on all types of information terminals, receiving downlink information and returning data. The active end provides the modulated light source of the entire system. The uplink relies on the downlink modulated light signal as the carrier wave and uses the reverse modulator to carry the secondary modulation to the downlink light signal. The uplink data are then modulated to the reflected light signal and reflected back to the active end. The secondary modulated light signal is processed at the active end, and the uplink signal is extracted.

LED Optical Design

The optical design is divided into primary and secondary optical designs. The primary optical design includes LED packaging and LED photoelectric components to determine the LED light angle, light intensity, light flux, light intensity

Table 33.5 Domestic and foreign research results in visible light communication

Year	Organization	Transmitter	Modulation	Equilibrium	Receiving end	Transmission rate	Distance
2008 [160]	Oxford University	White LED	OOK	Pre/Post	Sensors	75 Mbit/s	2.15 m
2008 [161]	Oxford University	White LED	OOK	Pre	PIN	80 Mbit/s	10 cm
2009 [162]	HHI, Germany	White LED	DMT	Post	PIN	230 Mbit/s	70 cm
2010 [163]	HHI, Germany	White LED	DMT	Post	APD	513 Mbit/s	30 cm
2011	Tripp College	RGB LED	DMT	Post	APD	3.4 Gbit/s	2 m
2012 [165]	Oxford University	RGB LED	MIMO	Pre/Post	PIN	1 Gbit/s	1 m
2013 [166]	Taiwan Jiaotong University	RGB LED	Cap	Pre/Post	APD	3.22 Gbit/s	25 cm
2014 [167]	Fudan University	RGB LED	SC	Post	PIN	4.22 Gbit/s	1 cm
2014 [168]	Italian SSSUP	RGBY LED	DMT	Pre/Post	PIN	5.6 Gbit/s	1.5 m
2015 [169]	Xi'an University of Technology	RGB LED	16-PPM	Post	PIN	1.5 Mbit/s	3 m
2015 [170]	Fudan University	RGB LED	Cap	Pre/Post	PIN	8 Gbit/s	1 m
2016 [171]	Oxford University	RGB LED	DMT	Pre/Post	PIN	10.4 Gbit/s	1.5 m
2017 [172]	Xi'an University of Technology	RGB LED	DQPSK	Pre/Post	PIN	5 Mbit/s	3 m
2017 [173]	University of Edinburgh	Micro-LED	OFDM	Pre/Post	PIN	10.72 Gbit/s	1 m
2019 [174]	Fudan University	White LED	DMT	Pre/Post	PIN	2.5 Gbit/s	1.7 m

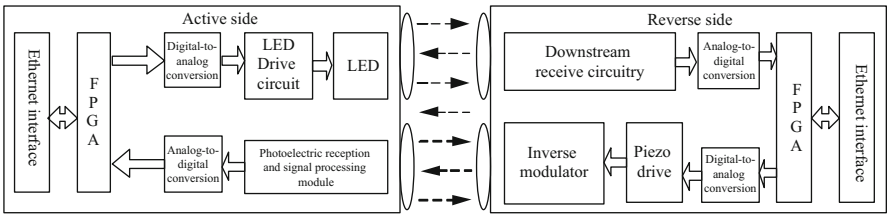


Fig. 33.23 Overall system architecture [181]

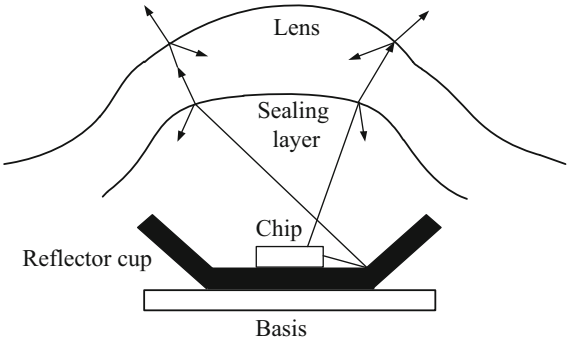


Fig. 33.24 Primary optical design diagram using a CREE-q5 flashlight [181]

distribution, and color temperature range and distribution. Figure 33.24 shows the primary optical design structure. This structure is designed to improve the light extraction rate so that the distribution of the half-light intensity angle is 120°.

To ensure the stability of indoor optical communication, the optical power on the receiving surface needs to be evenly distributed, and the secondary optical design adopts an optimized optical lens design algorithm. A VLC transmitter antenna [175] with uniform illuminance distribution, no inter-cell interference, and low design complexity can be realized.

The secondary optical design involves changing the optical performance of the beam after it passes through a lens and before passing through an optical lens. Therefore, the secondary optical design mainly optimizes the design of the LED lighting. The secondary optical light distribution design includes various elements, such as an external reflection cup, multiple optical lenses, and aspherical light surfaces, making it possible to respond to different scenarios.

The principle is shown in Fig. 33.25a, where the LED, after a single optical adjustment, diffuses light through a lens to a wider angle, mainly solving the lighting problem. In Fig. 33.25b, in contrast, the focusing problem is mainly solved by focusing the wide-angle light into a narrower light through a combination of mirrors and lenses or a single lens.

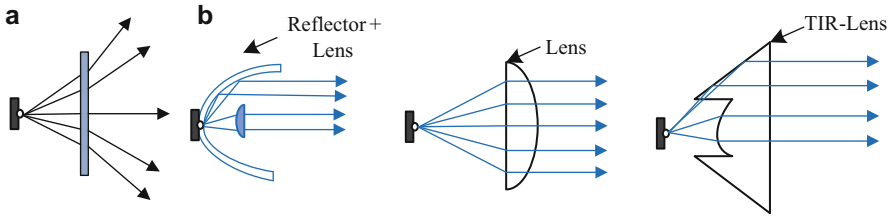


Fig. 33.25 Focused and astigmatic LED optical designs [174]

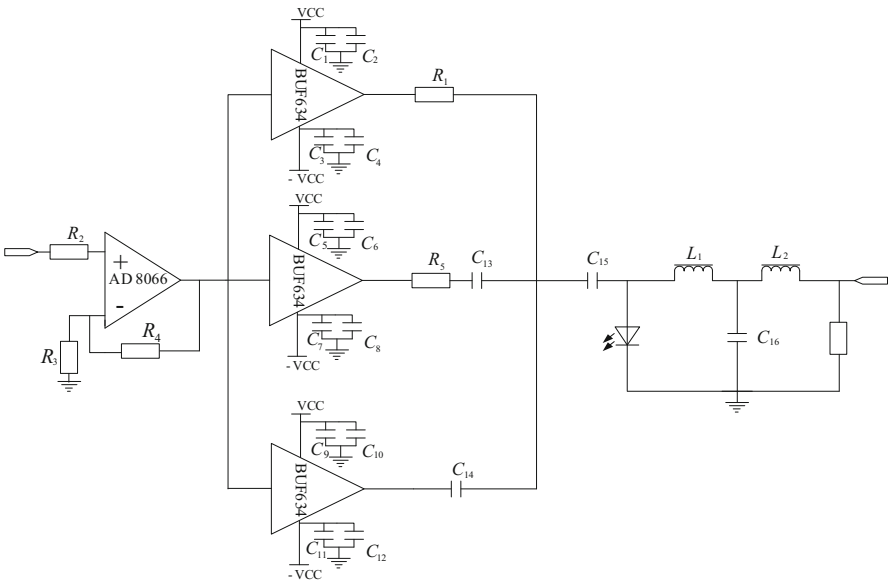


Fig. 33.26 LED driver circuit schematic [175]

LED Driver Circuit

The main function of the LED drive circuit is to convert the input voltage signal into a current signal and drive the LED to generate a modulated light-wave signal. The LED driver circuit consists of a DC bias circuit, preequalization circuit, and bias-T circuit [175]. The function of the DC bias circuit is to provide a stable one-way current to the LED and to make the LED work in a linear range.

The role of the preequalization circuit is to increase the modulation bandwidth of the LEDs by compensating for the fading component of the signal and reducing the signal distortion caused by the insufficient response strength of LEDs in the high-frequency band [176]. The main function of the bias-T circuit is to load the AC modulation component on the bias DC. The overall structure of the LED driver circuit is shown in Fig. 33.26.

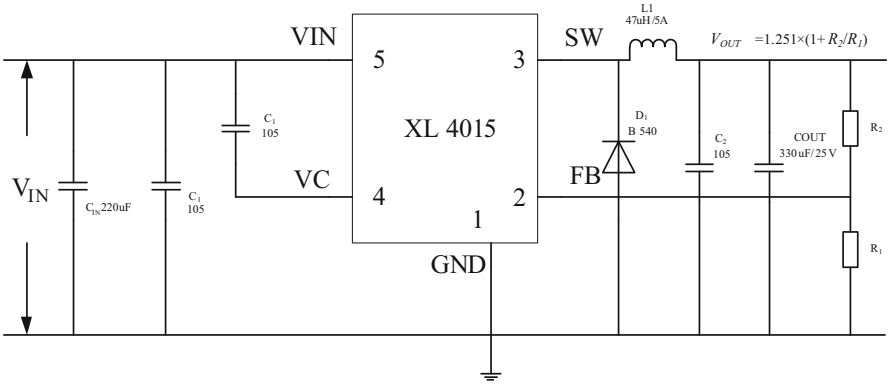


Fig. 33.27 DC bias circuit [175]

The LED driver requires a DC input; therefore, the LED modulation driver circuit biases the current source design requirements for a constant current source circuit. Figure 33.27 shows a bias constant current source circuit. XL4015 is a buck DC/DC converter with a fixed switching frequency of 180 kHz, which can provide up to a 5 A current output. The highest conversion efficiency is 93%. XL4015 has high integration, fewer peripheral devices, and flexible applications. The bias circuit design uses the output voltage and constant current output mode of the XL4015.

The modulation bandwidth of commercial high-power white LEDs is typically low, and the drive circuit plays an important role in extending their modulation bandwidth. A preequalizer can be used to increase the LED modulation bandwidth. When a white LED is used in a high-frequency circuit, its inductance characteristic exhibits a resonance effect with the capacitor in the circuit when the LED is working in the linear range.

The principle of the preequalization circuit is that the LED light source is equivalent to a low-pass filter with a cutoff frequency of 3 MHz in the driving circuit, while the preequalization circuit is equivalent to a band-pass filter with a pass frequency of 3 MHz or more. The central frequency of the resonance is controlled by changing the size of the capacitance in the circuit, extending the circuit in the high-frequency direction through the frequency range of the signal [177]. Considering that LED modulation requires a larger current, the buffer module in the pre-equalization circuit requires a higher current output capacity. The circuit is shown in Fig. 33.28.

The bias-T circuit is a three-port module that sets the DC bias points for certain electronic components without disturbing other electronic components [178]. Because of the unidirectional conductivity of the LED, the input modulation signal can only be a positive-amplitude voltage signal. Thus, it is necessary to transfer the AC signal produced by the modulator to a positive amplitude to be output from the LED.

The output signal power of the modulation circuit is very small, and its dynamic range exceeds the linear operating range of the LED. For LEDs to stably emit

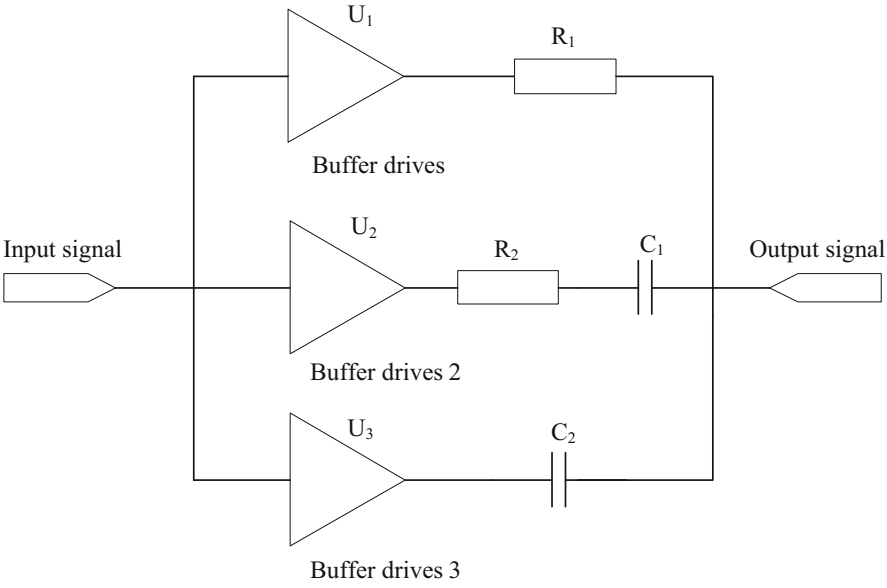
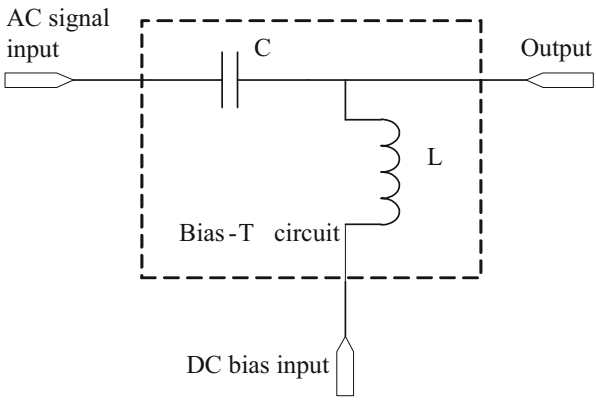


Fig. 33.28 Preequalizer circuit [175]

Fig. 33.29 Bias-T schematic [176]



modulated light signals, it is necessary to set a bias current to stabilize the bias of the LEDs and to superimpose the modulated current signal onto the bias current signal. The modulated signal is stabilized in the linear working range of the LED to obtain a good modulation effect.

Figure 33.29 shows a schematic diagram of the bias-T circuit. The DC bias is input through the inductor, the AC signal is input through the capacitor, and the two signals are superposed to the output through the output port.

Experimental Verification

Overall Uplink Design

In an optical full-duplex communication system, the uplink is a reverse communication link based on a reverse modulator. To realize a reverse communication link, we must first understand the characteristics of the reverse modulation device. There are many types of inverse modulators. Typical optical inverse modulators are divided into two types: the cube corner reflector and the "cat's eye" structure echo modulator [179].

A corner-reflecting prism is an optical device composed of three highly reflective planes placed perpendicular to each other. It is characterized by the fact that when the light beam is incident on any plane, the reflected light returns along the original incident side. Figure 33.30a, b show the attenuating reflection reverse modulator and externally modulated reverse modulator, respectively.

A reflector is placed on the focal plane of an optical lens, and the light rays are concentrated at the focal point of the lens after entering the lens. Owing to the reversibility of the light path, the reflector on the focal plane reflects the light beam back along the original path; this phenomenon is called the cat's eye effect. According to the difference in the modulation principle, the inverse modulation with a cat's eye structure can be divided into two types: defocus type and non-defocus type [180]. Figure 33.31a, b show electro-absorption-type and defocus-type inverse modulators, respectively.

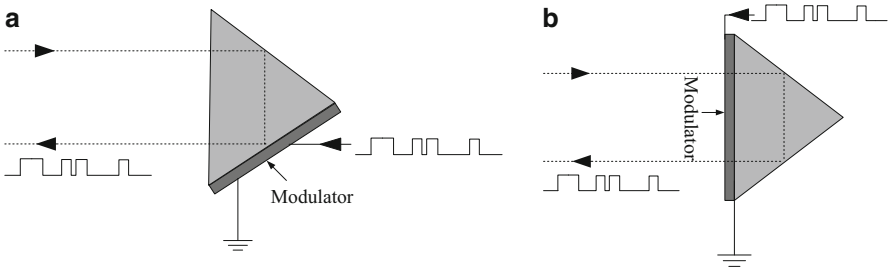


Fig. 33.30 Schematic diagram of the conic prism-type inverse modulator [179] (**a**) Attenuating reflection reverse modulator (**b**) Externally modulated reverse modulator

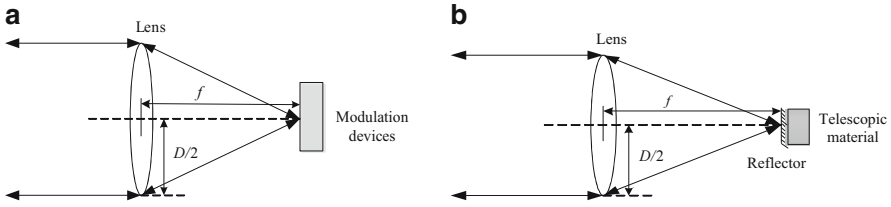


Fig. 33.31 Schematic diagrams of cat's eye reverse modulator models [180]. (**a**) Non-defocused inverse modulator (**b**) Defocused inverse modulator

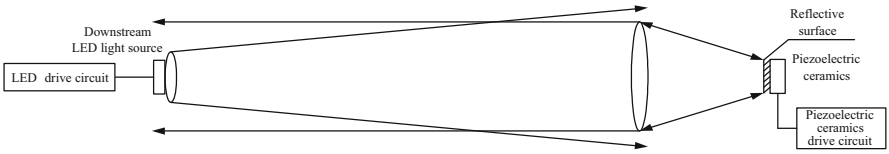


Fig. 33.32 Reverse communication system architecture diagram [181]

The uplink is a reverse communication link from the reverse end to the active end. The core of the reverse communication link is the cat's-eye reverse modulation system. The out-of-focus cat's-eye inverse modulator is composed of an optical lens, piezoelectric ceramic chip, modulation drive circuit, and echo-receiving circuit. Figure 33.32 shows the reverse communication system structure diagram.

When the system operates, the active end first sends the modulated light beam to the end of the reverse modulator. Because of the large divergence angle of the white LED beam, the receiving focusing lens and the cat's-eye reverse modulator are placed side by side at the reverse end and are located in the effective lighting area of the LED light source.

At the reverse end, the optical signal of the downlink is modulated twice by the reverse modulator, and the uplink information is modulated onto the downlink and reflected back to the active end. The active optical receiver receives the returned optical signal, separates the uplink modulation signal from it, and sends it to the demodulator for demodulation to realize a duplex communication link.

System Experiment

The Xi'an University of Technology experimental system was built, and the described full-duplex communication system was tested [178]. To verify the principle of single-light source full-duplex communication, a single-light source full-duplex communication experiment was carried out on the developed visible light full-duplex communication system.

The downlink adopted the differential quadrature phase-shift keying (DQPSK) modulation format; the carrier rate was set to 5.0 MHz, and the data rate was 5.0 Mbit/s. The uplink adopted the 4-way frequency-shift keying modulation format (4FSK) and used four carriers of different frequencies, f1, f2, f3, and f4, to represent different symbol combinations: 00 corresponds to f1, 01 to f2, 10 to f3, and 11 to f4. The carrier rate for f1 = 4 kHz, f2 = 8 kHz, f3 = 16 kHz, and f4 = 20 kHz. The data rate was 2.0 kbit/s.

During the experiment, the transmitted signal was collected, as shown in Figs. 33.33, 33.34, and 33.35. Figure 33.33 shows the measured figure of the signal sent by the active end and received by the reverse end in the downlink. Figure 33.34 shows the signal and its power spectrum after secondary modulation at the reverse end. Figure 33.35 shows a 4FSK modulated signal in the uplink.

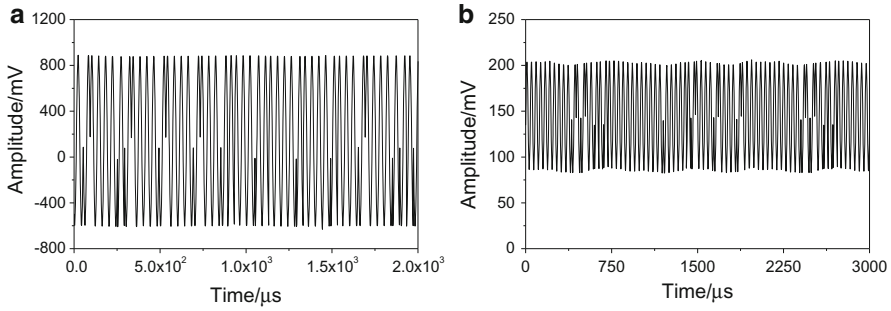


Fig. 33.33 Downlink communication signal [178]. (**a**) Active side sends the signal (**b**) Reverse side receives the signal

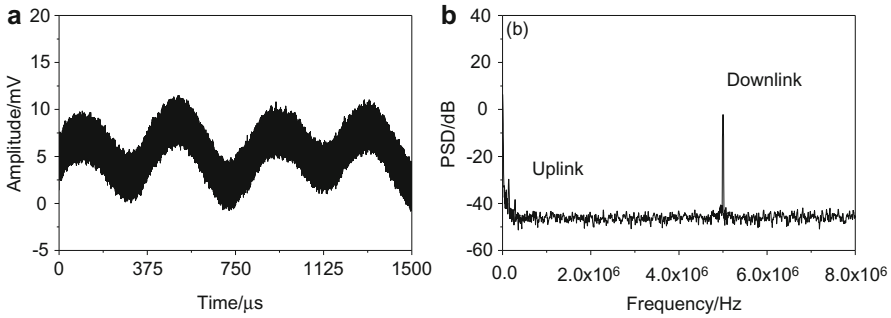


Fig. 33.34 Signal and power spectrum after secondary modulation [178]. (**a**) Signal after secondary modulation (**b**) Signal power spectrum

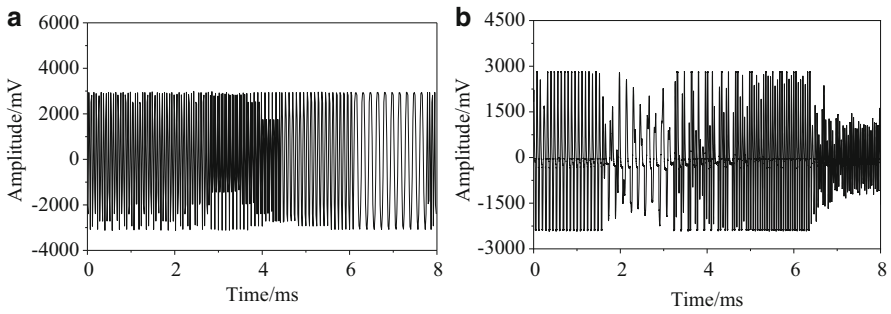


Fig. 33.35 Uplink Communication Link Signal [178]. (**a**) Reverse end transmits signal (**b**) Signal extracted by the active end

It can be seen from the graph that the received signal amplitude of the downlink is stable and clear, the visible light full-duplex communication system with an LED as a light source can meet the needs of reverse demodulation, and the two links can communicate without the influence of intersymbol crosstalk.

During the entire experiment, the downlink video communication was good, and there was no frame loss. The uplink data transmission involves normal, uplink, and downlink communication between each link without interference. It can be seen from the experimental results that full-duplex communication between the uplink and downlink can be realized independently and without interference by the two modulations of the unidirectional light source. The effectiveness of the inverse modulation technique for solving the uplink problem in indoor visible light communication was verified.

The experimental results show that the platform can achieve stable full-duplex communication with an active-end-to-reverse data transfer rate of 5.0 Mbit/s and a reverse-end-to-active transmission rate of 2.0 kbit/s over a communication distance of 3.0 m.

Summary

LEDs are widely used semiconductor light sources. They can be used for optical communication, with typical wavelengths of 780 nm, 850 nm, 1300 nm, etc. The LED manufacturing cost is low and reliable, the drive circuit is simple, and the temperature, emission power, and drive current have a good linear relationship (unless near saturation).

However, LED light sources have wide spectral lines (more than 20 nm, high dispersion) and divergent beams. They have low coupling efficiency with fiber, or need special lenses for coupling, and are not suitable for single-mode fiber applications with small core diameters. Therefore, they are only used for low-speed ($<$100 Mbit/s), short-distance ($<$100 m) communication.

At present, three regions—North America, Europe, and Asia—are the center of the global industrial pattern; each has a strong advantage in a different area. The developed economies, represented by the United States, Japan, and Germany, have the manufacturing capacity of the entire LED lighting industry chain and have an advantage in the high-end market, with advanced equipment levels and a relatively complete industrial chain system; however, owing to the high cost of labor, logistics, and supply-chain constraints, they do not scale well.

In emerging Asian market countries, India, Vietnam, and other countries represent the late development of the LED lighting industry. Although they have a certain product manufacturing capacity, their industrial chain is not perfect, it faces complex political and economic situations, and the social efficiency is low. In the short-term, it is difficult for these countries to be competitive.

China is the world's largest manufacturer and exporter of LED lighting products; it is extremely competitive in terms of production, research and development, and sales.

LED Industry Opportunities

Promote LED Lighting Products and Provide National Policy Support

To improve energy efficiency, protect the environment, and combat global climate change, the United States, European Union, Japan, and other developed countries/regions have introduced bans on traditional incandescent lamps and provided support to promote the development of LED lighting industry policies.

LED Lighting Technology Continues to Develop and Prices Continue to Decline

The rapid development of the LED lighting industry is determined by the progress of technology. In the lighting industry, LED lighting products have replaced energy-saving lamps. The progress in LED lighting technology is mainly reflected in two aspects.

The first is the continuous improvement of light efficiency. According to China's National LED Engineering Research and Development (R&D) and Industry Alliance, in 2018, the light efficiency of industrial high-power white LEDs exceeded 180 lm/mW, LED indoor lamps exceeded 100 lm/mW, and outdoor lamps exceeded 130 lm/mW.

The second is the continuous decline in cost. Owing to the improvement in upstream chip technology, the increase in production capacity, and the optimization of the production process, the cost of the main raw materials has gradually decreased, which reduces the production cost of LED lamps. The price of a 60-watt LED bulb was about $45 to $50 in January 2011; by July 2018, the price had fallen to less than $10 per bulb, and by May 2019, it had fallen further to $7. LED lighting is starting to approach the price of traditional energy-saving lamps.

LED Lighting Products Continue to Expand Their Application Fields

With the rapid development of LED lighting technology, LED lighting products' stability, service life, intelligence, light efficiency, and other performance indicators are becoming increasingly mature, to meet the market demand. LED lighting products not only meet the needs of general lighting. The market scope continues to extend and expand, according to different application scenarios, derived from a variety of application fields.

Before the 1990s, the LED lighting industry was in the early stage of development, with a single product color and higher prices. The application market was mainly driven by signal lights, indicator lights, and other similar demands. Then, the LED lighting industry entered a rapid development stage. The development of blue and green light technology and the decline in product prices rapidly promoted LED lighting in the commercial field and became the main driving force for the development of the LED lighting market at that time.

With the development of technology and the decrease in price, LED lighting products are widely used in home lighting, industrial lighting, etc. In the future, with the development of landscape lighting, automotive lighting, and other subindustries, the LED lighting market demand will continue to maintain rapid growth.

China Has a Complete LED Lighting Industry Chain and the Industry Is in a Mature Period of Development

China's LED lighting industry started in the late 1960s. After nearly 50 years of development, the lighting industry has made remarkable achievements, forming a complete industrial chain, including LED epitaxial chip production, preparation, packaging, and product application.

In addition, China attaches great importance to the research and development of LED lighting technology. It has launched green lighting and LED projects and supports R&D and industrialization of LED lighting technology in 10 key energy-saving projects, high-tech demonstration projects, enterprise technology upgrading and structural adjustment projects, and 863 new materials. Therefore, it has a good R&D basis and advantages in the integrated application of downstream LED lighting.

China's LED Industry Standards and Development Continue to Improve

In recent years, under the guidance of government departments, such as the Ministry of National Development and Reform Commission, the Standardization Administration of China, the Ministry of Science and Technology, the Ministry of Industry and Information Technology, and with the promotion of relevant industry organizations and bodies, the testing and certification of LED lighting standards in China has made phased progress.

According to statistics from the National Standard Information Public Service Platform, as of May 31, 2021, China had issued 187 current national standards, 174 industry standards, and 247 local standards related to LED lighting. It actively carries out group standard pilot work and actively participates in international ISO/TC274 standardization.

In addition, LED lighting products are increasingly used in a wide range of applications, such as outdoor areas, stadiums, and other downstream applications, and relevant management standards have been gradually established. In addition, the Ministry of Railways, Ministry of Transport, and Ministry of Public Security have established corresponding LED lighting application product standards. The establishment of these industry standards and norms will help promote the orderly and standardized development of the LED lighting industry.

LED Industry Challenges

The Core Technology Is Relatively Weak

At present, a large number of LED lighting products in China are used for export; however, the developed markets in Europe and North America are different from the domestic market. They have strict access thresholds, developed trust management mechanisms, and through high-tech regulations, high safety standards, and other means have set market access conditions for LED lighting products.

However, the core technologies of LED lighting enterprises in China have been weak for a long time in terms of high-end materials, optical design, power management, and heat dissipation technology. LED lighting enterprises have no competitive advantage in technology, which has become a major obstacle for LED lighting products being exported to developed countries/regions, such as Europe and the United States.

The Brand Influence of Domestic LED Lighting Enterprises Is Weak

The global LED lighting industry has formed a three-legged industrial distribution and competition pattern, dominated by the United States, Asia, and Europe. Japan, America, and Germany have the leading industries, followed by China, Taiwan, and Korea. They are followed by the echelon distribution of mainland Chinese countries and regions, such as Malaysia.

International giants, such as Xinnufei and Osram, occupy a large market share in the high-end product market and have their own characteristics in product and market aspects. They form a complete LED lighting industry chain and take advantage of its innovative advantages in the field of new products and technologies. They are mainly engaged in the production of high value-added products, master high-end materials, optical design, power management, cooling technology, and other key core technologies. However, domestic LED lighting enterprises have weak brand promotion and value-adding abilities or neglect brand construction in favor of sales scales.

The Homogeneous Competition Phenomenon Is Serious

Because the LED lighting market is in the high-speed growth stage, some small and medium-sized enterprises in the industry invest less in the research and development of products and technologies and occupy the market through using relatively low prices. This leads to the LED lighting industry having a certain degree of disorderly competition and product homogeneity. These problems not only affect the profit of LED lighting enterprises, but are also unfavorable to the health and sustainable development of the LED lighting industry.

References

1. Nair G B, Swart H C, Dhoble S J. A review on the advancements in phosphor-converted light emitting diodes (pc-LEDs): Phosphor synthesis, device fabrication and characterization [J]. Progress in Materials Science, 2020, 109(3): 1–36.
2. Song J, Lee H, Jeong E G, et al. Organic light-emitting diodes: pushing toward the limits and beyond [J]. Advanced Materials, 2020, 32(35): 1–19.
3. Feng Fangping, Liu Yi, Chen Baixing. Analysis of global LED patent time situation [J]. Guangdong Science and Technology, 2010, 19(17):38–43.
4. Zhang Fengching. Prospect of light-emitting diodes (LEDs)—LEDs will eventually replace incandescent lamps and fluorescent lamps [J]. Vacuum electronics, 2003, 1(3):74–74.
5. Krames M R, Shchekin O B, Mueller-Mach R, et al. Status and future of high-power light-emitting diodes for solid-state lighting [J]. Journal of display technology, 2007, 3(2): 160–175.

6. Wen Qidong. Export report of China's lighting industry in 2021[J]. Chinese lighting appliances, 2022, 1(3):1–12
7. Zhang Zhongtai, Zhang Junying. Inorganic photoluminescent materials and their applications [M]. Beijing: Chemical Industry Press, 2011: 100–111
8. Zhang Ning. Homogenization method of LED surface light source [J]. informatization research, 2012, 38(3): 46–49.
9. Xiao Zhiguo Semiconductor lighting luminescent materials and their applications [M]. Beijing: Chemical Industry Press, 2014: 30–38
10. Tamulaitis G, Duchovskis P, Bliznikas Z, et al. High-power light-emitting diode based facility for plant cultivation [J]. Journal of Physics D: Applied Physics, 2005, 38(17): 3182–3187.
11. Kuo Y K, Tsai M C, Yen S H, et al. Effect of p-type last barrier on efficiency droop of blue InGaN light-emitting diodes [J]. IEEE Journal of Quantum Electronics, 2010, 46(8): 1214–1223.
12. Tongjun Y, Matuo T, Suto K, et al. Green light emissions from GaP-AlxGa1− xP double heterostructures [J]. Journal of electronic materials, 1998, 27(9): 1053–1058.
13. Schnitzer I, Yablonovitch E, Caneau C, et al. Ultrahigh spontaneous emission quantum efficiency, 99.7% internally and 72% externally, from AlGaAs/GaAs/AlGaAs double heterostructures [J]. Applied physics letters, 1993, 62(2): 131–133.
14. Wang Y, Zhen C, Gong H, et al.Measurement of the specific contact resistance of Au/Ti/p-diamond using transmission line model [J]. Acta Physica Sinica, 2000, 49(7): 1348–1351.
15. Chi H, Du Y, Li G. Analysis of InGaN Multiple-Quantum-Well Photoelectric Device of Visible Light Communication [J]. Journal of Nanoelectronics and Optoelectronics, 2020, 15(7): 909–918.
16. Dumke W P. Spontaneous radiative recombination in semiconductors [J]. Physical Review, 1957, 105(1): 139–144.
17. Weisbuch C, Piccardo M, Martinelli L, et al. The efficiency challenge of nitride light-emitting diodes for lighting [J]. physica Status Solidi (a), 2015, 212(5): 899–913.
18. Hu Guoyong, Chen Changying, Chen Zhenqiang. Research on the use of white LED lighting source as indoor wireless communication [J]. Optical communication technology, 2006, 30(7): 46–48.
19. Xia Pengkun, Cheng Qijia. Effect of PN junction doping concentration on depletion layer width, built-in electric field and built-in potential [J]. University Physics, 2015, 34(6): 54–56.
20. Preparation of ZnO/Si heterojunction and its photoelectric characteristics [M]. Chengdu: University of Electronic Science and Technology of China Press, 2017: 70–78.
21. Guo Haozhong, Lai Fangyi, Guo Shouyi. LED principle and application [M]. Beijing: Chemical Industry Press, 2013: 30–35.
22. Komine T, Haruyama S, Nakagawa M. A study of shadowing on indoor visible-light wireless communication utilizing plural white LED lightings [J]. Wireless Personal Communications, 2005, 34(1): 211–229.
23. Lee J S, Su Y W, Shen C C. A comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi [C]. 2007, Annual Conference of the IEEE Industrial Electronics Society: 46–51.
24. Ding Jupeng, Huang Zhitong. Research direction and technical optimization of optical and wireless communication [J]. the world of digital communication, 2009 30(8): 66–68.
25. Fang Zhilie. History, status and prospect of light-emitting diode materials and devices [J]. physics, 2003, 32(05): 01–06.
26. Liu Tingzhang. Modular design technology of LED lighting drive power supply [M]. Beijing: China Machine Press, 2018, 7: 15–56
27. Holonyak Jr N, Bevacqua S F. Coherent (visible) light emission from GaN (As1− x P x) junctions [J]. Applied Physics Letters, 1962, 1(4): 82–83.
28. Nakamura S, Senoh M S M, Mukai T M T. Highly p-typed Mg-doped GaN films grown with GaN buffer layers [J]. Japanese journal of applied physics, 1991, 30(10A): L1708–L1711.

29. Ilegems M, Dingle R. Luminescence of Be-and Mg-doped GaN [J]. Journal of Applied Physics, 1973, 44(9): 4234–4235.
30. Amano H, Sawaki N, Akasaki I, et al. Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer [J]. Applied Physics Letters, 1986, 48(5): 353–355.
31. Nakamura S N S. GaN growth using GaN buffer layer [J]. Japanese Journal of Applied Physics, 1991, 30(10A): L1705–L1707.
32. Höfler G E, Vanderwater D A, DeFevere D C, et al. Wafer bonding of 50-mm diameter GaP to AlGaInP-GaP light-emitting diode wafers [J]. Applied Physics Letters, 1996, 69(6): 803–805.
33. Strielkowski W, Veinbender T, Tvaronavičienė M, et al. Economic efficiency and energy security of smart cities [J]. Economic research-Ekonomska istraživanja, 2020, 33(1): 788–803.
34. Hwang J I, Hashimoto R, Saito S, et al. Development of InGaN-based red LED grown on (0001) polar surface [J]. Applied Physics Express, 2014, 7(7): 1–4.
35. Hwang D, Mughal A, Pynn C D, et al. Sustained high external quantum efficiency in ultrasmall blue III–nitride micro-LEDs [J]. Applied Physics Express, 2017, 10(3): 1–4.
36. Bi Z, Lu T, Colvin J, et al. Realization of ultrahigh quality InGaN platelets to be used as relaxed templates for red micro-LEDs [J]. ACS applied materials & interfaces, 2020, 12(15): 17845–17851.
37. Mukai T, Narimatsu H, Nakamura S. Amber InGaN-based light-emitting diodes operable at high ambient temperatures [J]. Japanese journal of applied physics, 1998, 37(5A): L479–L481.
38. Mukai T, Yamada M, Nakamura S. Characteristics of InGaN-based UV/blue/green/amber/red light-emitting diodes [J]. Japanese Journal of Applied Physics, 1999, 38(7R): 3976–3981.
39. Mukai T. Recent progress in group-III nitride light-emitting diodes [J]. IEEE Journal of selected topics in quantum electronics, 2002, 8(2): 264–270.
40. Ohkawa K, Watanabe T, Sakamoto M, et al. 740-nm emission from InGaN-based LEDs on c-plane sapphire substrates by MOVPE [J]. Journal of crystal growth, 2012, 343(1): 13–18.
41. Yang G F, Zhang Q, Wang J, et al. InGaN/GaN multiple quantum wells on selectively grown GaN microfacets and the applications for phosphor-free white light-emitting diodes [J]. Reviews in Physics, 2016, 1(1): 101–121.
42. Li P, Li H, Li Z, et al. Strong carrier localization effect in carrier dynamics of 585 nm InGaN amber light-emitting diodes [J]. Journal of Applied Physics, 2015, 117(7): 1–5.
43. Iida D, Lu S, Hirahara S, et al. Enhanced light output power of InGaN-based amber LEDs by strain-compensating AlN/AlGaN barriers [J]. Journal of Crystal Growth, 2016, 1(448): 105–108.
44. Iida D, Niwa K, Kamiyama S, et al. Demonstration of InGaN-based orange LEDs with hybrid multiple-quantum-wells structure [J]. Applied Physics Express, 2016, 9(11): 1–4.
45. Lee K, Lee H, Lee C R, et al. Yellow-red light-emitting diodes using periodic Ga-flow interruption during deposition of InGaN well [J]. Optics Express, 2017, 25(13): 15152–15160.
46. Mitchell B, Dierolf V, Gregorkiewicz T, et al. Perspective: Toward efficient GaN-based red light emitting diodes using europium doping [J]. Journal of Applied Physics, 2018, 123(16): 1–12.
47. Jiang F, Zhang J, Xu L, et al. Efficient InGaN-based yellow-light-emitting diodes [J]. Photonics Research, 2019, 7(2): 144–148.
48. Dussaigne A, Barbier F, Damilano B, et al. Full InGaN red light emitting diodes [J]. Journal of Applied Physics, 2020, 128(13): 1–9.
49. Iida D, Zhuang Z, Kirilenko P, et al. 633-nm InGaN-based red LEDs grown on thick underlying GaN layers with reduced in-plane residual stress [J]. Applied Physics Letters, 2020, 116(16): 1–5.
50. Iida D, Zhuang Z, Kirilenko P, et al. Demonstration of low forward voltage InGaN-based red LEDs [J]. Applied Physics Express, 2020, 13(3): 1–5.
51. Pasayat S S, Gupta C, Wong M S, et al. Demonstration of ultra-small (< 10 μm) 632 nm red InGaN micro-LEDs with useful on-wafer external quantum efficiency ($> 0.2\%$) for mini-displays [J]. Applied Physics Express, 2020, 14(1): 1–4.

52. Zhuang Z, Iida D, Ohkawa K. Investigation of InGaN-based red/green micro-light-emitting diodes [J]. Optics Letters, 2021, 46(8): 1912–1915.
53. Thomas D G, Hopfield J J. Isoelectronic traps due to nitrogen in gallium phosphide [J]. Physical Review, 1966, 150(2): 856–680.
54. Damilano B, Gil B. Yellow–red emission from (Ga, In) N heterostructures [J]. Journal of Physics D: Applied Physics, 2015, 48(40):1–28.
55. Groves W O, Herzog A H, Craford M G. The Effect of Nitrogen Doping on GaAs1− x P x Electroluminescent Diodes [J]. Applied Physics Letters, 1971, 19(6): 184–186.
56. Kuo C P, Fletcher R M, Osentowski T D, et al. High performance AlGaInP visible light-emitting diodes [J]. Applied Physics Letters, 1990, 57(27): 2937–2939.
57. Sato H, Chung R B, Hirasawa H, et al. Optical properties of yellow light-emitting diodes grown on semipolar (11 2 2) bulk GaN substrates [J]. Applied Physics Letters, 2008, 92(22): 1–3.
58. Saito S, Hashimoto R, Hwang J, et al. InGaN light-emitting diodes on c-face sapphire substrates in green gap spectral range [J]. Applied Physics Express, 2013, 6(11): 1–4.
59. Gessmann T, Schubert E F. High-efficiency AlGaInP light-emitting diodes for solid-state lighting applications [J]. Journal of applied physics, 2004, 95(5): 2203–2218.
60. Jiang Fengyi, Liu Junlin, Zhang Jianjian, et al. New materials, new devices and new equipment for semiconductor yellow light-emitting diodes [J]. Acta Physica Sinica, 2019, 68(16): 331–339.
61. Zhao X, Tang B, Gong L, et al. Rational construction of staggered InGaN quantum wells for efficient yellow light-emitting diodes [J]. Applied Physics Letters, 2021, 118(18): 1–6.
62. Kang J M, Kim J W, Choi J H, et al. Life-time estimation of high-power blue light-emitting diode chips [J]. Microelectronics Reliability, 2009, 49(11): 1231–1235.
63. Lester S D, Ludowise M J, Killeen K P, et al. High-efficiency InGaN MQW blue and green LEDs [J]. Journal of crystal growth, 1998, 189(1): 786–789.
64. Jiang Y, Li Y, Deng Z, et al. Realization of high-luminous-efficiency InGaN light-emitting diodes in the "green gap" range [J]. Scientific Reports, 2015, 5(11): 1–7.
65. Li P, Zhao Y, Li H, et al. Highly efficient InGaN green mini-size flip-chip light-emitting diodes with AlGaN insertion layer [J]. Nanotechnology, 2019, 30(9): 1–5.
66. Shao X, Yan D, Lu H, et al. Efficiency droop behavior of GaN-based light emitting diodes under reverse-current and high-temperature stress [J]. Solid-state electronics, 2011, 57(1): 9–13.
67. Wang L, Jin J, Mi C, et al. A review on experimental measurements for understanding efficiency droop in InGaN-based light-emitting diodes [J]. Materials, 2017, 10(11): 1–27.
68. Tan Z K, Moghaddam R S, Lai M L, et al. Bright light-emitting diodes based on organometal halide perovskite [J]. Nature nanotechnology, 2014, 9(9): 687–692.
69. Veldhuis S A, Boix P P, Yantara N, et al. Perovskite materials for light-emitting diodes and lasers [J]. Advanced materials, 2016, 28(32): 6804–6834.
70. Li G, Tan Z K, Di D, et al. Efficient light-emitting diodes based on nanocrystalline perovskite in a dielectric polymer matrix [J]. Nano letters, 2015, 15(4): 2640–2644.
71. Wang J, Wang N, Jin Y, et al. Interfacial control toward efficient and low-voltage perovskite light-emitting diodes [J]. Advanced Materials, 2015, 27(14): 2311–2318.
72. Heo J M, Cho H, Lee S C, et al. Bright Lead-Free Inorganic CsSnBr3 Perovskite Light-Emitting Diodes [J]. ACS Energy Letters, 2022, 7(8): 2807–2815.
73. Li J, Shan X, Bade S G R, et al. Single-layer halide perovskite light-emitting diodes with sub-band gap turn-on voltage and high brightness [J]. The journal of physical chemistry letters, 2016, 7(20): 4059–4066.
74. Xiao Z, Kerner R A, Zhao L, et al. Efficient perovskite light-emitting diodes featuring nanometre-sized crystallites [J]. Nature Photonics, 2017, 11(2): 108–115.
75. Zhang L, Yang X, Jiang Q, et al. Ultra-bright and highly efficient inorganic based perovskite light-emitting diodes [J]. Nature communications, 2017, 8(1): 1–8.

76. Yang X, Zhang X, Deng J, et al. Efficient green light-emitting diodes based on quasi-two-dimensional composition and phase engineered perovskite with surface passivation [J]. Nature communications, 2018, 9(1): 1–8.
77. Li Xiuqing. Research Status of GaN Blue-UV-UV Optoelectronic Materials and Devices [J]. Semiconductor intelligence, 2000, 37(2): 58–64.
78. Li Jinmin. Progress in the development of SiC materials and devices [J]. Acta Physica Sinica, 2000, 29(8): 481–487.
79. Yoshida S, Misawa S, Gonda S. Improvements on the electrical and luminescent properties of reactive molecular beam epitaxially grown GaN films by using AlN-coated sapphire substrates [J]. Applied Physics Letters, 1983, 42(5): 427–429.
80. Nakamura S, Harada Y, Seno M. Novel metalorganic chemical vapor deposition system for GaN growth [J]. Applied physics letters, 1991, 58(18): 2021–2027.
81. Qin Z X, Chen Z Z, Tong Y Z, et al. Study of Ti/Au, Ti/Al/Au, and Ti/Al/Ni/Au ohmic contacts to n-GaN [J]. Applied Physics A, 2004, 78(5): 729–731.
82. Nakamura S, Mukai T, Senoh M. Candela-class high-brightness InGaN/AlGaN double-heterostructure blue-light-emitting diodes [J]. Applied Physics Letters, 1994, 64(13): 1687–1689.
83. Jin S X, Li J, Li J Z, et al. GaN microdisk light emitting diodes [J]. Applied Physics Letters, 2000, 76(5): 631–633.
84. Choi H W, Jeon C W, Dawson M D. Fabrication of matrix-addressable micro-LED arrays based on a novel etch technique [J]. Journal of crystal growth, 2004, 268(3): 527–530.
85. Gong Z, Gu E, Jin S R, et al. Efficient flip-chip InGaN micro-pixellated light-emitting diode arrays: promising candidates for micro-displays and colour conversion [J]. Journal of Physics D: Applied Physics, 2008, 41(9):1–6.
86. Tian P, McKendry J J D, Gong Z, et al. Characteristics and applications of micro-pixelated GaN-based light emitting diodes on Si substrates [J]. Journal of Applied Physics, 2014, 115(3): 1–4.
87. Zhang L, Ou F, Chong W C, et al. Wafer-scale monolithic hybrid integration of Si-based IC and III–V epi-layers—A mass manufacturable approach for active matrix micro-LED micro--displays [J]. Journal of the Society for Information Display, 2018, 26(3): 137–145.
88. Mo X F, Xu W Z, Lu H, et al. Fabrication and Characterization of a GaN-Based Micro-LED Array [J]. Chin. Phys. Lett, 2017, 34(11): 98–101.
89. Qi L, Zhang X, Chong W C, et al. 848 ppi high-brightness active-matrix micro-LED micro-display using GaN-on-Si epi-wafers towards mass production [J]. Optics Express, 2021, 29(7): 10580–10591.
90. Zheng Daishun, Qian Keyuan. Research progress of power type white LED [J]. China Lighting Appliances,2006, 1(3): 1–7.
91. Grundmann M J, Mishra U K. Multi-color light emitting diode using polarization-induced tunnel junctions [J]. physica status solidi c, 2007, 4(7): 2830–2833.
92. Li T, Luo J, Honda Z, et al. Sintering condition and optical properties of Zn3V2O8 phosphor [J]. Adv Mater Phys Chem, 2012, 2(1): 173–177.
93. Sato Y, Ishizaki S, Murakami Y, et al. Some properties of group-III nitride thin films directly grown on non-single-crystalline substrates by using a molecular beam epitaxy apparatus [J]. Journal of Modern Physics, 2015, 6(9): 1289–1297.
94. Maruska H P, Stevenson D A, Pankove J I. Violet luminescence of Mg-doped GaN [J]. Applied Physics Letters, 1973, 22(6): 303–305.
95. Welker T. Recent developments on phosphors for fluorescent lamps and cathode-ray tubes [J]. Journal of luminescence, 1991, 48(49): 49–56.
96. Cho J, Park J H, Kim J K, et al. White light-emitting diodes: history, progress, and future [J]. Laser & photonics reviews, 2017, 11(2): 1–19.
97. Xiao D, Kim K W, Bedair SM, et al. Design of white light-emitting diodes using InGaN/AlInGaN quantum-well structures [J]. Applied Physics Letters, 2004, 84(5):672–674.

98. Fu D, Jing X. Chromaticity Simulation for the Influence of the Bule Phosphor Spectrum to FED Properties [J]. Chinese Journal of Luminescence, 2004, 25(3): 320–328.
99. Schubert E F. Light-emitting diodes [M]. Rensselaer: E. Fred Schubert, 2018: 20–40.
100. Misra A, Kumar P, Kamalasanan M N, et al. White organic LEDs and their recent advancements [J]. Semiconductor science and Technology, 2006, 21(7): 35–35.
101. Templier F. GaN-based emissive microdisplays: a very promising technology for compact, ultra-high brightness display systems [J]. Journal of the Society for Information Display, 2016, 24(11): 669–675.
102. Schlotter P, Schmidt R, Schneider J. Luminescence conversion of blue light emitting diodes [J]. Applied Physics A: Materials Science & Processing, 1997, 64(4): 15–36.
103. Moreno I, Avendaño-Alejo M, Tzonchev R I. Designing light-emitting diode arrays for uniform near-field irradiance [J]. Applied optics, 2006, 45(10): 2265–2272.
104. Narukawa Y. White-light LEDS [J]. Optics and photonics news, 2004, 15(4): 24–29.
105. Narukawa Y, Ichikawa M, Sanga D, et al. White light emitting diodes with super-high luminous efficacy [J]. Journal of physics D: Applied physics, 2010, 43(35): 1–6.
106. Evgenyevna Z O, Vasilyevna S L, Nikolaevich A S. Features of the Methodological Approach to the Evaluation of the LED Lighting Efficiency in the Industrial Premises [J]. Journal of fundamental and Applied Sciences, 2017, 9(7S): 856–862.
107. Bergh A, Craford G, Duggal A, et al. The promise and challenge of solid-state lighting [J]. Physics today, 2001, 54(12): 42–47.
108. Yu Bo. LEDinside:2013 "Top Ten Conjectures" of China's LED Industry [J]. China Electronics Business Information (Basic Electronics), 2013, 1(3):53–57.
109. Goofy. In 2022, China's lighting field will move forward in innovative applications [J]. Journal of Lighting Engineering, 2022, 33(1): 6–6.
110. Yan Jianchang, Sun Lili, Ran Junxue, Wang Junxi. Research and development hotspots of ultraviolet light-emitting diode materials and devices [J]. Journal of Lighting Engineering, 2018, 29(6): 10–10.
111. Detchprohm T, Hiramatsu K, Amano H, et al. Hydride vapor phase epitaxial growth of a high quality GaN film using a ZnO buffer layer [J]. Applied physics letters, 1992, 61(22): 2688–2690.
112. Han J, Crawford M H, Shul R J, et al. AlGaN/GaN quantum well ultraviolet light emitting diodes [J]. Applied physics letters, 1998, 73(12): 1688–1690.
113. Nishida T, Kobayashi N. 346 nm Emission from AlGaN Multi-Quantum-Well Light Emitting Diode [J]. physica status solidi (a), 1999, 176(1): 45–48.
114. Kinoshita A, Hirayama H, Ainoya M, et al. Room-temperature operation at 333 nm of Al 0.03 Ga 0.97 N/Al 0.25 Ga 0.75 N quantum-well light-emitting diodes with Mg-doped superlattice layers [J]. Applied Physics Letters, 2000, 77(2): 175–177.
115. Nishida T, Saito H, Kobayashi N. Efficient and high-power AlGaN-based ultraviolet light-emitting diode grown on bulk GaN [J]. Applied Physics Letters, 2001, 79(6): 711–712.
116. Taniyasu Y, Kasu M, Makimoto T. An aluminium nitride light-emitting diode with a wavelength of 210 nanometres [J]. nature, 2006, 441(7091): 325–333.
117. Hirayama H, Fujikawa S, Noguchi N, et al. 222–282 nm AlGaN and InAlGaN-based deep-UV LEDs fabricated on high-quality AlN on sapphire [J]. physica status solidi (a), 2009, 206(6): 1176–1182.
118. Fujioka A, Misaki T, Murayama T, et al. Improvement in output power of 280-nm deep ultraviolet light-emitting diode by using AlGaN multi quantum wells [J]. Applied physics express, 2010, 3(4): 1–3.
119. Kinoshita T, Obata T, Nagashima T, et al. Performance and reliability of deep-ultraviolet light-emitting diodes fabricated on AlN substrates prepared by hydride vapor phase epitaxy [J]. Applied Physics Express, 2013, 6(9): 1–3.
120. Hirayama H, Maeda N, Fujikawa S, et al. Recent progress and future prospects of AlGaN-based high-efficiency deep-ultraviolet light-emitting diodes [J]. Japanese Journal of Applied Physics, 2014, 53(10): 1–10.

121. Hideki H. Growth Techniques of AlN/AlGaN and Development of High-Efficiency Deep-Ultraviolet Light-Emitting Diodes [M]. New York: Springer Intemational Publishing Switzerand, 2016: 75–113.
122. Shatalov M, Jain R, Dobrinsky A, et al. High-efficiency UV LEDs on sapphire [A]. Gallium Nitride Materials and Devices X [C]. California, United States: SPIE, 2015, 9363: 203–209.
123. Inoue S, Tamari N, Taniguchi M. 150 mW deep-ultraviolet light-emitting diodes with large-area AlN nanophotonic light-extraction structure emitting at 265 nm [J]. Applied Physics Letters, 2017, 110(14): 1–5.
124. Zhang Z H, Huang Chen S W, Chu C, et al. Nearly efficiency-droop-free AlGaN-based ultraviolet light-emitting diodes with a specifically designed superlattice p-type electron blocking layer for high Mg doping efficiency [J]. Nanoscale Research Letters, 2018, 13(1): 1–7.
125. Lee J W, Ha G, Park J, et al. AlGaN deep-ultraviolet light-emitting diodes with localized surface plasmon resonance by a high-density array of 40 nm Al nanoparticles [J]. ACS applied materials & interfaces, 2020, 12(32): 36339–36346.
126. Li Y, Xing Z, Zheng Y, et al. High-efficiency near-UV light-emitting diodes on Si substrates with InGaN/GaN/AlGaN/GaN multiple quantum wells [J]. Journal of Materials Chemistry C, 2020, 8(3): 883–888.
127. Ponce F A, Bour D P. Nitride-based semiconductors for blue and green light-emitting devices [J]. Nature, 1997, 386(6623): 351–359.
128. Amano H, Kito M, Hiramatsu K, et al. P-type conduction in Mg-doped GaN treated with low-energy electron beam irradiation (LEEBI)[J]. Japanese journal of applied physics, 1989, 28(12A): L2112–L2114.
129. Nakamura S, Mukai T M T, Senoh M S M. High-power GaN pn junction blue-light-emitting diodes [J]. Japanese Journal of Applied Physics, 1991, 30(12A): L1998–L2001.
130. Sugawara H, Ishikawa M, Hatakoshi G. High-efficiency InGaAlP/GaAs visible light-emitting diodes [J]. Applied physics letters, 1991, 58(10): 1010–1012.
131. Huang K H, Yu J G, Kuo C P, et al. Twofold efficiency improvement in high performance AlGaInP light-emitting diodes in the 555–620 nm spectral region using a thick GaP window layer [J]. Applied physics letters, 1992, 61(9): 1045–1047.
132. Nakamura S, Senoh M S M, Mukai T M T. P-GaN/N-InGaN/N-GaN double-heterostructure blue-light-emitting diodes [J]. Japanese Journal of Applied Physics, 1993, 32(1A): L8–L11
133. Nakamura S, Mukai T, Senoh M. High-brightness InGaN/AlGaN double-heterostructure blue-green-light-emitting diodes [J]. Journal of Applied Physics, 1994, 76(12): 8189–8191.
134. Nakamura S, Senoh M, Iwasa N, et al. High-power InGaN single-quantum-well-structure blue and violet light-emitting diodes [J]. Applied Physics Letters, 1995, 67(13): 1868–1870.
135. Wang Xiaohui, Liu Xianglin, Lu Dacheng, et al. InGaN/GaN single quantum well green light-emitting diode [J].Journal of Semiconductors,2000, 1(7): 726–733.
136. Modak P, D'Hondt M, Delbeke D, et al. AlGaInP microcavity light-emitting diodes at 650 nm on Ge substrates [J]. IEEE Photonics Technology Letters, 2000, 12(8): 957–959.
137. Lai W C, Chang S J, Yokoyam M, et al. InGaN-AlInGaN multiquantum-well LEDs [J]. IEEE Photonics Technology Letters, 2001, 13(6): 559–561.
138. Tadatomo K, Okagawa H, Ohuchi Y, et al. High output power InGaN ultraviolet light-emitting diodes fabricated on patterned substrates using metalorganic vapor phase epitaxy [J]. Japanese Journal of Applied Physics, 2001, 40(6B): 121–129.
139. Koike M, Shibata N, Kato H, et al. Development of high efficiency GaN-based multiquantum-well light-emitting diodes and their applications [J]. IEEE Journal of selected topics in quantum electronics, 2002, 8(2): 271–277.
140. Shchekin O B, Epler J E, Trottier T A, et al. High performance thin-film flip-chip InGaN–GaN light-emitting diodes [J]. Applied Physics Letters, 2006, 89(7): 1–3.
141. Zhou S, Liu X, Yan H, et al. Highly efficient GaN-based high-power flip-chip light-emitting diodes [J]. Optics express, 2019, 27(12): A669–A692.
142. Chi Nan, Shi Meng, Niu Wenqing. LiFi: Development Status and Prospect of Visible Light Communication Technology [J]. Journal of Lighting Engineering, 2019, 30(1): 1–9.

143. Gfeller F R, Bapst U. Wireless in-house data communication via diffuse infrared radiation [J]. Proceedings of the IEEE, 1979, 67(11): 1474–1486.
144. Barry J R, Kahn J M, Krause W J, et al. Simulation of multipath impulse response for indoor wireless optical channels [J]. IEEE, 1993, 11(3): 367–379.
145. Tanaka Y, Haruyama S, Nakagawa M. Wireless optical transmissions with white colored LED for wireless home links [A]. 11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications. PIMRC 2000. Proceedings [C]. London, UK: IEEE, 2000, 2: 1325–1329.
146. Tanaka Y, Komine T, Haruyama S, et al. Indoor visible light data transmission system utilizing white LED lights [J]. IEICE transactions on communications, 2003, 86(8): 2440–2454.
147. Ding Deqiang, Ke Xizheng. Design and simulation of wireless LAN system based on visible light communication [J]. Journal of Xi'an University of Technology, 2007, 1(1): 29–32.
148. Arikan E. A performance comparison of polar codes and Reed-Muller codes [J]. IEEE Communications Letters, 2008, 12(6): 447–449.
149. Komine T, Lee J H, Haruyama S, et al. Adaptive equalization system for visible light wireless communication utilizing multiple white LED lighting equipment [J]. IEEE Transactions on Wireless Communications, 2009, 8(6): 2892–2900.
150. Ding Deqiang, Ke Xizheng. Research on a universal white LED mathematical luminescence model [J]. Acta Optica Sinica, 2010, 30(9): 2536–2540.
151. Bandara K D, Niroopan P, Yeon-Ho C. Improved indoor visible light communication with PAM and RLS decision feedback equalizer [J]. Iete Journal Of Research, 2013, 59(6): 672–678.
152. Ren G, He S, Yang Y. An improved recursive channel model for indoor visible light communication systems [J]. Information Technology Journal, 2013, 12(6): 1245–1250.
153. Li X, Bamiedakis N, Guo X, et al. Wireless visible light communications employing feed-forward pre-equalization and PAM-4 modulation [J]. Journal of Lightwave Technology, 2016, 34(8): 2049–2055.
154. Fang J, Che Z, Jiang Z L, et al. An efficient flicker-free FEC coding scheme for dimmable visible light communication based on polar codes [J]. IEEE Photonics Journal, 2017, 9(3): 1–10.
155. Wang H, Kim S. Dimming control systems with polar codes in visible light communication [J]. IEEE Photonics Technology Letters, 2017, 29(19): 1651–1654.
156. Yao S, Zhang X, Qian H, et al. Joint dimming and data transmission optimization for multi-user visible light communication system [J]. IEEE Access, 2017, 5(1): 5455–5462.
157. Bian R, Tavakkolnia I, Haas H. 15.73 Gb/s visible light communication with off-the-shelf LEDs [J]. Journal of Lightwave Technology, 2019, 37(10): 2418–2428.
158. Jani M, Garg P, Gupta A. Performance analysis of a mixed cooperative PLC–VLC system for indoor communication systems [J]. IEEE Systems Journal, 2019, 14(1): 469–476.
159. Wang L, Wang X, Kang J, et al. A 75-Mb/s RGB PAM-4 visible light communication transceiver system with pre-and post-equalization [J]. Journal of Lightwave Technology, 2020, 39(5): 1381–1390.
160. Le Minh H, O'Brien D, Faulkner G, et al. High-speed visible light communications using multiple-resonant equalization [J]. IEEE photonics technology letters, 2008, 20(14): 1243–1245.
161. Le Minh H, O'Brien D, Faulkner G, et al. 80 Mbit/s visible light communications using pre-equalized white LED [C]//2008 34th European Conference on Optical Communication. IEEE, 2008: 1–2.
162. Vucic J, Kottke C, Nerreter S, et al. White light wireless transmission at 200 Mb/s net data rate by use of discrete-multitone modulation [J]. IEEE Photonics Technology Letters, 2009, 21(20): 1511–1513.
163. Vučić J, Kottke C, Nerreter S, et al. 513 Mbit/s visible light communications link based on DMT-modulation of a white LED [J]. Journal of lightwave technology, 2010, 28(24): 3512–3518.

164. Cossu G, Khalid A M, Choudhury P, et al. 3.4 Gbit/s visible optical wireless transmission based on RGB LED [J]. Optics express, 2012, 20(26): B501–B506.
165. Azhar A H, Tran T A, O'Brien D. A gigabit/s indoor wireless transmission using MIMO-OFDM visible-light communications [J]. IEEE photonics technology letters, 2012, 25(2): 171–174.
166. Wu F M, Lin C T, Wei C C, et al. 3.22-Gb/s WDM visible light communication of a single RGB LED employing carrier-less amplitude and phase modulation [C]//2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC). IEEE, 2013: 1–3.
167. Wang Y, Huang X, Zhang J, et al. Enhanced performance of visible light communication employing 512-QAM N-SC-FDE and DD-LMS [J]. Optics express, 2014, 22(13): 15328–15334.
168. Cossu G, Wajahat A, Corsini R, et al. 5.6 Gbit/s downlink and 1.5 Gbit/s uplink optical wireless transmission at indoor distances ($\geq$ 1.5 m)[C]//2014 The European Conference on Optical Communication (ECOC). IEEE, 2014: 1–3.
169. Qang Ye, Ke Xizheng. Multidimensional Coding in Visible Light Communication [J]. Chinese laser, 2015, 42(02): 138–144.
170. Wang Y, Tao L, Huang X, et al. 8-Gb/s RGBY LED-based WDM VLC system employing high-order CAP modulation and hybrid post equalizer [J]. IEEE photonics journal, 2015, 7(6): 1–7.
171. Chun H, Rajbhandari S, Faulkner G, et al. LED based wavelength division multiplexed 10 Gb/s visible light communications [J]. Journal of lightwave technology, 2016, 34(13): 3047–3052.
172. Ke Xizheng, Zhang Huan. Indoor visible uplink communication system based on reverse modulation [J]. Advances in lasers and optoelectronics, 2018, 55(2): 115–126.
173. Islim M S, Ferreira R X, He X, et al. Towards 10 Gb/s orthogonal frequency division multiplexing-based visible light communication using a GaN violet micro-LED [J]. Photonics Research, 2017, 5(2): A35–A43.
174. Wang C, Li G, Ha Y, et al. A 2.5 Gb/s Real-Time Visible-Light Communication System B ased on Phosphorescent White LED [C]. 2019, 7th International Conference on Information, Communication and Networks (ICICN): 140–145.
175. Wang Yongbo. Research on indoor visible light wireless communication system [D]. Zhejiang: Zhejiang University, 2008. 30–31.
176. Li Rongling. Visible light communication system based on white LED [D]. Shanghai: Fudan University, 2014. 11–14.
177. Song Xiaoqing, Wei Youcai, Zhao Zixu, et al. Research Progress of Hardware Pre-equalization Technology for High-speed Visible Light Communication [J]. Advances in lasers and optoelectronics, 2017, 54(8): 26–36.
178. Hao L M, O'Brien D, Faulkner G, et al. 80 Mbit/s Visible Light Communications using pre-equalized white LED [C]. 2008, Brussels, Belgium: European Conference on Optical Communication: 1–2.
179. Li Zhan, Peng Renjun, WU Jian, et al. Optical reverse modulation technology [J]. Advances in lasers and optoelectronics, 2007, 44(11): 47–51.
180. Qiu Hao, Wang Jingyuan, XU Zhiyong, et al. Reverse modulation optical communication technology [J]. Military communication technology, 2015, 36 (2): 29–33.
181. Zhang Huan. Design of visible light full-duplex communication system [D]. Xi'an: Xi'an University of Technology, 2018: 80–83.

Principle and Research Progress of Visible Light Communication

34

Contents

Introduction 1308
Visible LED Communications 1308
LED Light-Emitting Principle 1308
LED Communications 1310
Current Domestic and Foreign Research Status 1311
Current Foreign Research Status 1311
Current Status of Domestic Research 1314
Key Technologies in LED Visible Light Communication Systems 1318
Link Technology 1318
Reverse Modulation Techniques 1319
Equalization Technology 1320
Information Modulation Technology 1322
Bandwidth Extension Technology for White LED Light Sources 1323
Optical Camera Communication Technology 1324
Indoor Visible Light Communication LED Layouts 1325
Indoor Visible Light System Modeling 1325
Light Source Layout Design 1326
Visible Light Positioning 1327
Visible Light Positioning Technology 1327
Positioning Enablement Program 1328
Visible Light Networking 1328
Networking Overview 1328
Heterogeneous Networking 1332
Application Scenarios 1334
Smart Home 1334
Health Care 1335
Intelligent Transportation 1335
Aviation Applications 1335
Ship Communication 1335
Intelligent Industrial Lighting 1336
Museum Environment Monitoring 1336
Library Navigation 1336
Underwater Applications 1336
Emergency Applications 1337

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_34

Development Trends 1337
Summary and Outlook 1338
References 1339

Abstract

Visible light-emitting diode (LED) communication combines lighting and communication and has the advantages of high efficiency and energy saving, which are highly valued. This paper starts by elaborating on the LED light-emitting principle, introduces the progress of visible LED communication research, and analyzes some key technologies in the communication system. For the indoor application of white LED communication, the layout of LED light sources and visible light positioning are discussed, as well as several common visible light heterogeneous networking schemes. Finally, the future development direction of visible light communication is presented.

Keywords

LED · Visible light communication · Indoor LED layout · Visible light positioning · Heterogeneous network

Introduction

With the development of communication technologies, the traditional airwave spectrum is becoming increasingly crowded. Visible light communication (VLC) can make up for the lack of spectrum resources. White LEDs have high brightness, low power consumption, no pollution, long life, high emitted power, and other characteristics and are gradually replacing incandescent lamps to become the mainstream devices in the field of lighting [1].

In addition, VLC provides a communication function, making it a new wireless communication information carrier. Because of the many advantages of visible light communication, we conducted an in-depth study of the communication performance, including experimental simulations and other aspects, to explore the key technology of visible light communication. Future development trends and application prospects are presented to promote the development of visible light communication in all aspects.

Visible LED Communications

LED Light-Emitting Principle

LEDs, or light-emitting diodes, are mainly composed of LED chips, cathode rods, anode rods, lead frames, wedge-shaped brackets, etc., and then encapsulated in transparent epoxy resin [2]. Light-emitting diodes have unidirectional conductivity,

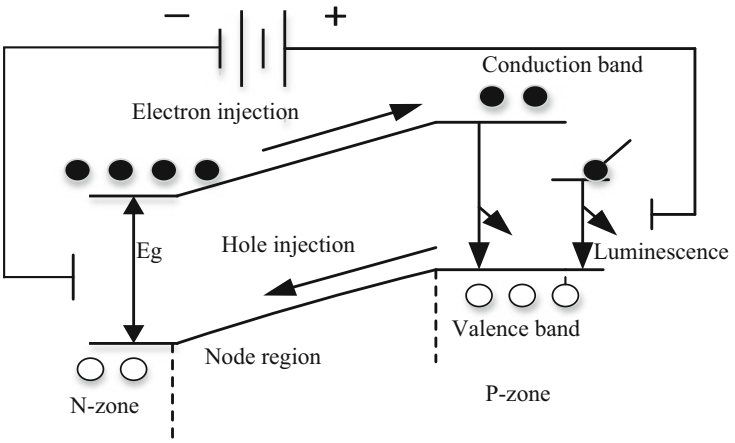


Fig. 34.1 LED light-emitting principle diagram [5]

the core of which is a semiconductor-doped P-N junction. The luminescence mechanism is compound luminescence using carrier radiation from the P-N junction, as shown in Fig. 34.1 [3].

When coupled with a forward bias voltage, the LED internal current flows through electron (N-type carriers) and hole (P-type carriers) composites in the light form to release excess energy by converting electrical energy into light energy [4]. The emitted light wave is the energy released by a high energy-level particle dropping to a lower energy level, and its frequency is $\nu = E_g/h$. Each particle is non-coherent, and the phase, polarization direction, and propagation direction of the emitted light wave are not the same. Therefore, LED luminescence is a spontaneous radiation process [5].

White light is a type of composite light. Typical white light LEDs are implemented using a multicolor LED combination method and phosphor-coating light transformation method, as described below [6, 7]:

Yellow phosphor is mixed on a blue LED chip to emit white light; as the white light generation must excite the fluorescent material, the modulation bandwidth is only a few tens of MHz, which does not meet the requirements of high-speed visible light communication systems.

An ultraviolet or violet LED is used as the light source and mixed with red, blue, and green phosphors to form white light according to the proportion. This method has a low luminous efficiency, and the modulation bandwidth has difficulty meeting the requirements of high-speed visible light communication systems.

A red, blue, and green three-color LED chip is packaged together; the three lights together emit white light. Although this method has a high production cost, it has a high diffractive color and the modulation response speed is very fast; thus, it is very suitable for high-speed visible light communication systems.

LED Communications

Visible light communication (VLC) is a optical wireless communication technology that uses light with a wavelength of 380 nm ~ 780 nm as an information carrier to transmit information in free space. The information is modulated by the LED transmitting light carrier, photodetectors and other photoelectric conversion devices are used to receive the light carrier signals, and the receiving end analyzes the received light intensity signals by judging and restoring the original information. This combines visible light communication with LED lighting, constructing LED lighting, and communicating dual-use base station lights [8].

The LED visible light communication system structure shown in Fig. 34.2 mainly consists of the transmitting and receiving parts. The optical signal transmitting part includes input and processing circuits (used to convert the signal source signals into electrical signals conducive to transmission in the channel) and a visible light drive modulation circuit (used to modulate the changes in electrical signals into changes in the intensity of the light carrier).

The optical signal receiving section consists of an optical system (for optimal reception of the signal source), a photodetector and preamplification circuit (for converting the optical signal to an electrical signal), and a signal processing and output circuit (to convert the electrical signal into a signal that can be recognized by the terminal).

Table 34.1 compares several wireless communication methods, including LED visible light communication [5, 9].

It can be seen that LED visible light communication has the following advantages [10]:

The spectrum does not require authorization. This is an effective solution to the current situation of tight wireless spectrum resources.

It can be used for simultaneous lighting and communication. It can realize point-to-point terminal barrier-free communication and reduce signal loss under lighting conditions.

Communication is safe and reliable. Because visible light cannot penetrate walls and other obstacles, the security and privacy of information are guaranteed.

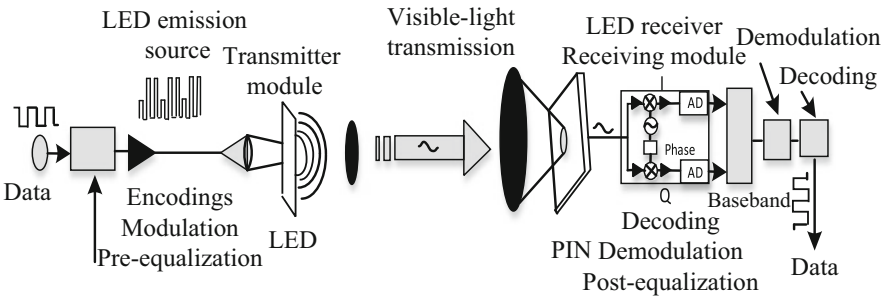


Fig. 34.2 Diagram of LED visible light communication system [8]

Table 34.1 Comparison of communication methods [5, 9]

Comparator	Radio	Infrared	Visible light
Is spectrum certification required?	Yes	No	No
Does it go through obstacles?	Yes	No	No
Presence of multipath fading	Yes	No	No
Presence of multipath dispersion	Yes	Yes	Yes
Is power limited?	Yes	Yes	No
Path loss	High	High	High
Main sources of noise	Other users	Backlighting	Backlighting
Shadow effect	Unaffected	Vulnerability to shelter	Mounting multiple LEDs eliminates shadow effects

It is safe for human eyes. Visible light LEDs are safe, radiation-free, and will not harm the human eye.

It has no electromagnetic interference. The signal transmission is not subject to other external interference and can be applied to the more sensitive environment of electromagnetic induction.

It has small shadow effect. Compared with infrared light, which is easily blocked and leads to a shadow effect, the shadow effect can be eliminated by installing multiple LED lights.

It has low construction and maintenance costs. The optical path can quickly form a temporary physical channel without the need to set up additional expensive base stations.

Current Domestic and Foreign Research Status

Current Foreign Research Status

In 2000, Tanaka et al. [11] from Keio University in Japan proposed a white LED communication system for wireless home links. Firstly, the light source was modeled as a Lambertian radiation model. Then, the direct view and non-direct view links were investigated, and multipath effects and intercode crosstalk were analyzed to avoid degrading the system performance. Numerical analyses and simulation studies confirmed that white light LEDs had both illumination and communication functions.

In 2001, Tanaka et al. [12] used a computer to conduct an analog simulation of the channel of a white LED indoor communication system, using on-off keying-return-to-zero (OOK-RZ) and orthogonal frequency-division multiplexing (OFDM)

modulation to address the effect of the optical range difference generated by multiple LED light sources.

In 2002, Fan and Komine et al. [13] of Nakagawa Research Laboratory investigated the effect of wall reflections on visible light communication systems. Their analyses revealed that the communication performance characteristics deteriorated at locations close to a wall, owing to increased interference from the reflected light signals.

In 2003, Komine et al. [14] proposed a wiring system that combined power line carrier communication and visible light communication, i.e., using existing wires for visible light communication. It remedied the visible light communication wiring problem of LEDs and also emitted LED light without demodulation, according to the signal wave transmitted by the power line. Its feasibility was proved through experimentation.

In 2004, Douseki et al. [15] introduced a self-powered optical wireless data transmission system that used solar cells and operated under white LED illumination. Using nonzeroed Manchester pulse digital optical modulation combined with an ultra-low-power CPU, a high-speed transmission of more than 100 kbit/s was successfully achieved over a 40-cm illumination distance.

The BER performance of the visible optical communication system is degraded due to intercode interference. In 2005, T. Komine et al. [16] proposed an adaptive equalization technique based on the minimum mean square error algorithm to effectively mitigate the intercode crosstalk.

In 2006, Sugiyatna et al. [17] proposed an inverted pulse position modulation and subcarrier inverted pulse position modulation technique, and the results showed that the subcarrier modulation can suppress the effect of background light.

In 2006, Afgani et al. [18] utilized extremely bright, white LEDs for optical wireless communication using OFDM. It was concluded that the high peak-to-average ratio property of OFDM could be constructively utilized for LED intensity modulation in visible light communication.

In 2007, Haruyama et al. [19] proposed two brightness control methods for lighting and visible light communication systems: pulse-width modulation (PWM) and changing the modulation depth. The results of the study showed that using PWM control controlled the brightness from 0% to 87.5%; when changing the modulation depth, the brightness could be controlled in the range of 0% to 100%, and the communication performance was better than that of PWM.

In 2008, Japanese scholars demonstrated an LED system using duplex communication at the Electronic High-Tech Exhibition. The signal transmission rate reached 100 Mbit/s with a communication distance of 20 cm between the transmitter and receiver [3].

In 2009, because the optical range difference between multiple light sources triggers intercode interference, which seriously degrades the system performance, Komine et al. [20] proposed a method that combined a judgment feedback equalizer with a least-mean-square algorithm. The results showed that the adaptive equalization system effectively mitigated the effect of shadows.

In 2009, Vucic et al. [21] from the Technical University of Berlin, Germany, achieved 200-Mbit/s visible light communication using 32 50-MHz subcarriers for discrete multitone (DMT) modulation.

In 2010, researchers in Heinrich Hertz's laboratory built a visible light communication system using ordinary commercial white LEDs and increased the communication rate to 513 Mbit/s using 128,100-MHz subcarriers, which was the world's leading level at that time [22].

In 2011, the IEEE promulgated the IEEE standard 802.15.7-2011 for short-range optical wireless communication using visible light. It was the first internationally accepted visible light communication standard and defined the physical and MAC layers for short-range visible light communication [23].

In 2012, Fujimoto et al. [24] completed experiments with a maximum transmission speed of 614 Mbit/s using a single commercially available red LED, for the first time. They employed pre- and post-equalization circuits and a simple RC network at both the optical transmitter and receiver ends using OOK modulation. They also confirmed the error-free operation of the proposed high-speed VLC system using an RGB-type white LED with a transmission speed of 603 Mbit/s.

In 2013, Bandara et al. [25] used multi-pulse amplitude modulation to significantly increase the data rate, while maintaining the LED switching rate limit. The receiver side was equalized using a recursive least-squares decision feedback equalizer to reduce the sequence length and improve throughput.

In 2014, Tsonev et al. [26] designed a VLC system using a single 50 μm gallium nitride LED using OFDM as a modulation scheme, which allows the limited modulation bandwidth of gallium nitride μLED devices to be fully utilized. Equalization and post-equalization techniques and adaptive data loading techniques were also successfully applied to achieve a VLC communication rate of 3Gb/s.

In 2015, Edirisinghe et al. [27] designed and implemented a bidirectional LED optical communication system. They conducted mathematical modeling and experimental evaluations of uplink and downlink crosstalk using infrared and white LEDs, respectively, and presented crosstalk cancelation methods.

In 2017, Miramirkhani et al. [28] proposed a visible light communication channel model for mobile users. Non-sequential ray tracing was used to determine the channel impact response at each point in the mobile user's trajectory, and then, the path loss and delay extensions were fitted to a curve as a function of distance. Experimental results showed that applying the adaptive scheme to a mobile channel allowed the transmission parameters to be adjusted according to changes in the channel environment.

In 2017, Islim et al. [29] achieved a data rate of 11.95 Gbit/s when considering only the nonlinear distortion of the LED device as the dominant noise source. A data rate of 7.91 Gbit/s was produced when all noise sources were present, using a violet micro-LED as the transmitter source for an OFDM-based VLC system.

In 2018, to address the problem of a high peak-to-average power ratio, owing to the limited dynamic range of LEDs, Ebrahimi et al. [30] proposed a hybrid OFDM-pulse-time modulation (PTM) scheme, which converted bipolar optical OFDM

samples into a pulse-width-modulated and pulse position-modulated digital PTM format for the intensity modulation of LEDs, which improved the bit error rate.

In 2019, Bian et al. [31] designed a four-color wavelength-division multiplexing system with DC-biased optical OFDM and optimized the experimental parameters. The data transmission rate reached 15.73 Gbit/s over a 1.6-m wireless link. In the same year, Ghassemlooy et al. [32] at Northumbria University, UK, proposed a relay-assisted scheme for visible light communication to enhance the stability of the system, which could help in the popularization and application of visible light communication technology.

In 2020, Mohapatra et al. [33] investigated a single-input single-output visible light communication system subjected to additive Gaussian white noise interference. A closed expression for the symbol error probability of the system was derived with complete knowledge of the channel and a least-squares estimation of the channel at the receiver.

In 2021, Raj et al. [34] developed an analytical model of a dimmable VLC system using channel equalization. It was analyzed and used to study the effect of dimming and the data rate on the error performance of variable switch keying and variable pulse position modulation schemes.

In 2021, Aravindan et al. [35] developed a VLC system for transmission using a 3-m indoor free-space direct vision link using a single phosphorescent white LED. It was proposed to overcome the bandwidth limitation of LEDs using dot pre-equalization circuits with nonzero switch keying to achieve data transmission rates up to 1.3 Gbit/s.

In 2022, Poulose [36] used the OptiSystem simulation software to investigate direct and non-direct path models for indoor VLC systems. The non-rectilinear link propagation model used a single LED as the transmitter and varied the system bit rate when the field of view varied.

In 2023, Agha et al. [37] investigated the order of multiple-input multiple-output (MIMO) arrays for an interference-free capacity system. The final simulation and theoretical results showed that the inter-cell interference in visible light communication was greatly mitigated using the MIMO system.

Current Status of Domestic Research

In 2006, Hu et al. [38] compared the characteristics of white LED and infrared wireless communication technologies and proposed a white LED indoor wireless communication technology. By analyzing the channel and link, the factors to be considered in designing an indoor optical wireless communication system were determined.

In 2006, Ding et al. [39] analyzed the structure of a VLC indoor system and further proposed that the optimal light source layout of the room, optimal number of LEDs, and selection of the receiver field of view (FOV) were key technologies to be considered for indoor VLC.

In 2008, Chen et al. [40] proposed a theoretical scheme for angle diversity optical receiver technology and gave the corresponding photodetector (PD) layout model for this scheme. It improved the signal-to-noise ratio of the communication system, as well as overcoming intersymbol interference.

In 2009, Cao et al. [41] analyzed the indoor channel characteristics in terms of both the background optical noise and the multipath effect and used angle diversity reception as a solution. The research results showed that the diversity reception technique effectively improved the system performance and laid a solid foundation for achieving stable communication in high-speed networks.

In 2010, Ding et al. [42] analyzed the multipath channel model of an indoor visible light communication system and proposed a new reflection model of the Lambert–Phong model after studying the reflection characteristics of the channel model. A Monte Carlo ray-tracing simulation was completed under the line of this model, and the results showed that the simulation calculation of the ray-tracing algorithm using the Lambert–Phong model had a linear relationship with the number of reflections, which reduced the computational complexity.

In 2012, Zhang et al. [43] proposed a multi-input single-output subcarrier-multiplexed VLC system. By analyzing the variation characteristics of the BER with the transmission distance, they concluded that the transmission was reliable when the communication distance was less than 140 cm.

In 2013, Chi [44] proposed a new visible light communication system with full-duplex bidirectional subcarrier multiplexing wave-division multiplexing (WDM). It used quadrature amplitude modulation–OFDM and pre- and post-equalization techniques to achieve a visible light communication rate of 225 Mbit/s for the uplink based on P-LEDs and 575 Mbit/s for the downlink based on RGB-LEDs.

In 2013, Li et al. [45] studied the optical–electrical–optical interconversion relationship between the transmitter/receiver and the transfer function of the generalized visible light link, combined with the time/frequency-domain response characteristics. They constructed a complete mathematical model of the visible light communication system and obtained the influence law of the signal-to-noise ratio and the bit error rate of the visible light link on the transmission rate. When the signal-to-noise ratio was fixed, the limiting transmission rate had an e exponential growth relationship with the BER; for a fixed BER level, the limiting transmission rate changed with the signal-to-noise ratio as a Boltzmann function.

In 2014, Wang et al. [46] proposed a parallel high-speed transmission scheme based on OOK modulation for a nonlinear low-communication channel model in VLC. The scheme used an interleaved-division multiple access technique to simultaneously transmit from multiple wick sources. Compared with the multicarrier DC-biased optical OFDM under the same conditions, this scheme had a lower bit error rate and better nonlinearity resistance.

In 2015, Guan et al. [47] proposed a dual-modulation technique for LED light sources based on chip-on-board packages. Specifically, a dual-modulation technique combining pulse amplitude modulation and pulse wavelength modulation was applied to modulate and dim the LED light source system, respectively. Simulation

results showed that the VLC system using the dual-modulation technique had a lower BER than the system using digital switch keying modulation.

In 2015, Kang et al. [48] proposed a multidimensional coding method to improve the spectrum utilization of visible light communication, built a visible light audio-transmission system based on RGB-LEDs, and experimentally validated it using 16-pulse position modulation for multicolor LEDs. The results showed that this coding method effectively utilized the spectrum resources and increased the channel capacity.

In 2016, Yang et al. [49] introduced a peak average-to-power ratio reduction technique, i.e., OFDM mu-law compression, into the VLC system. The required signal format for VLC transmission was achieved and a lower peak average-to-power ratio signal was obtained.

In 2017, Wang et al. [50] studied the effect of the visible light communication receiving plane tilt angle on the channel capacity. After analyzing the system model, the upper and lower limits of the VLC channel capacity were given. It was found that the channel capacity could be greatly improved by appropriately tilting the receiving plane.

In 2018, Ke et al. [51] designed an indoor visible light single light source full-duplex communication system with an out-of-focus catadioptric inverse modulation scheme to achieve mutually independent link communication. The experimental results yielded uplink and downlink communication rates of 2.0 kbit/s and 5.0 Mbit/s, respectively.

In 2019, Wei et al. [52] achieved a transmission rate of 40 Gbit/s in 2-m free space using polarization-multiplexed RGB laser diodes, using orthogonal frequency-division multiplexing.

In 2020, Wu et al. [53] proposed a scheme to introduce deep learning techniques into channel evaluation for VLC systems, using DC-biased OFDM as the fundamental and relying on deep neural network models of deep learning techniques to estimate the channel with low spectral occupancy. This model had better BER performance than the traditional least-mean-square error method and also reduced the computational complexity of the least-mean-square error algorithm.

In 2021, Yang et al. [54] designed an integrated power carrier and visible light communication system for sites with weak electromagnetic signals. After laboratory and tunnel field tests, it simultaneously supported five handheld terminals online for real-time positioning.

In 2021, Ke et al. [55] proposed the use of a double-Gaussian function superposition to represent the VLC background light noise model for vehicular networking in nighttime environments and conducted real measurements at night in different regions. The results showed that there was no significant difference in the probability density distribution under both rainy and sunny weather conditions. Moreover, the light intensity fluctuation of street lights was found to be less than that produced by other light sources.

In 2022, Li et al. [56] proposed a noise-suppressed three-layer hybrid optical OFDM VLC system, which was further combined with pulse-width modulation to control the dimming. This dimming scheme achieved a wide dimming range of 89.0% while maintaining a BER below 3.8×10^{-3}.

In 2023, Chen et al. [57] introduced a joint power and subchannel resource allocation algorithm into an OFDM multicolor visible optical communication system. The scheme sufficiently improved resource utilization and obtained a higher transmission rate.

In 2023, Yang [58] used an artificial neural network to estimate and suppress the nonlinear distortion of the signal at the receiving end. The results showed that the method was able to reduce the bit error rate in different channel environments and effectively improved the performance of the visible light communication system.

Since the concept of visible light communication appeared in 2000, scholars at home and abroad have deeply investigated each part of the system. They have proposed various schemes to increase the communication rate and improve communication performance. In addition, they have explored the theoretical model under certain communication environments to provide a solid theoretical foundation for practical applications. Research approaches include experiments and simulations, but are not limited to them. The research content has gradually changed from the communication field to cross-cutting research in other fields, such as deep learning and neural networks. The above research on communication rates is summarized in Table 34.2.

Table 34.2 Visible light communication achievements

Year	Authors	Modulation	Result
2004	Douseki et al. [15]	Nonzero Manchester pulse digital modulation	Successful high-speed transmission of more than 100 kbit/s at an illuminated distance of 40 cm
2009	Vucic et al. [21]	DMT	Achieved 200 Mbit/s visible light communication using 32 50-MHz subcarriers
2010	Vucic et al. [22]	DMT	Increased communication rate to 513 Mbit/s using 128,100-MHz subcarriers
2012	Fujimoto et al. [24]	OOK	First experiment using a single commercially available red LED with a transmission speed of 614 Mbit/s
2013	Chi et al. [44]	Orthogonal amplitude modulation–OFDM	P-LED-based uplink with 225-Mbit/s communication rate, RGB-LED downlink rate of 575 Mbit/s
2014	Tsonev et al. [26]	OFDM	3-Gbit/s VLC communication rate achieved using GaN μLED devices
2017	Islim et al. [29]	OFDM	Achieved 10.72-Gbit/s transmission rate at 1-m communication distance using pre- and post-equalization techniques
2018	Ke et al. [51]	Out-of-focus cat's eye reverse modulation	5.0-Mbit/s downlink communication rate, 2.0-kbit/s uplink communication rate in duplex communication
2019	Bian et al. [31]	DCO-OFDM	Accomplished 15.73-Gbit/s real-time transfer rate over 1.6-m VLC link
2019	Wei et al. [52]	OFDM	40 Gbit/s in 2-m free space
2021	Aravindan et al. [35]	NRZ-OOK	1.3-Gbit/s data rate in 3-m indoor free space

Key Technologies in LED Visible Light Communication Systems

Link Technology

The directionality of light in visible light communication, its susceptibility to occlusion, and the complexity and diversity of noise sources result in an optical link model that is vastly different from that of traditional wireless communication [59]. According to the chronological order of light transmission and reception, the visible light link consists of three parts, i.e., a visible light communication light source, visible light transmission, and visible light reception.

Uplinks and Downlinks

A visible light communication system has two links between the receiving and transmitting ends: the uplink and the downlink. In downlink communication, optical transmitters and optical receivers use non-directional transmissions with unrestricted pointing and range; in uplink communication, optical transmitters and receivers use directional transmissions with restricted pointing and range [60].

An uplink communication is achieved by direct alignment, and in the case of duplex communication, the light intensity at the transmitter end of the downlink is stronger and the radiation angle of the LED light source is large; consequently, it will cover the uplink channel and interfere with the optical signal at the receiver end of the uplink [61].

Therefore, a crucial and important issue in uplink realization lies in avoiding interference from downlinks with illumination [62]. Based on uplink communication technologies implemented using radio waves and wireless light waves, the following uplink schemes have been proposed for visible light communication [63, 64]:

Infrared light in the 780–950-nm band is used as the uplink to distinguish visible light from the downlink.

VLC and RF are combined, i.e., using VLC as the downlink and a radio-frequency (RF) system as the uplink.

A laser reflector is used to reflect some of the incident light back to the transmitting system and modulate the reflected light to achieve the uplink transmission.

A visible light uplink is used such as having the downlink communicate using white LEDs to transmit data. The uplink communication can also use visible light bands for data transmission. Time-division techniques are used to separate the information transmission time for the uplink and downlink.

A polarizing beam splitter is used to separate the transmitting light from the receiving light, thus eliminating the effect of the downlink light on the uplink light.

Direct Vision Links vs. on-Direct Vision Links

In applications, a line-of-sight (LOS) or non-line-of-sight (NLOS) link is determined, based on whether the link is obstructed by objects in the path [65]. The unobstructed linear propagation of optical signals between the transmitter and

receiver is a line-of-sight link, while the presence of obstructions and reflections resulting in nonlinear propagation produces a non-line-of-sight link.

In an indoor visible light communication system, the direct view link has no wall reflections between the LED light source and the receiving end. Its transmission path is shorter, which increases the system communication rate, and high-speed data communication can be achieved. The path loss is smaller, the background light interference is weaker, there is no multipath fading effect, and the utilization rate of the signal light source is high.

The transmission path of the non-direct view link is longer, and the signal power is reduced after the signal has been reflected by a wall, etc. The paths arriving at the receiver at different times cause intercode interference, which degrades the reception performance. However, the coverage range is slightly larger than that of the direct view link, and some shadow effects can be overcome [66, 67]. The two links are compared in Table 34.3.

Reverse Modulation Techniques

Reverse modulation technology can be used to solve the uplink problem of indoor VLC. The principle is to use the reverse modulator for modulation by the incident light, following the original path back to the transmitter. Thus, the communication process can be completed without needing an active light source.

The key technology to achieve dual full-time communication is inverse modulation technology, in which, the inverse modulator is the core device. Common inverse modulators include angular cone prism reflection types and cat's eye structure echo-modulation types [68]. Figure 34.3 shows the schematic diagram of a cat's eye structure modulator [67]. According to the cat's eye structure, the inverse modulator is divided into two types of modulators: non-defocused and defocused. The two modulator models are shown in Fig. 34.4a and b, respectively [68].

Inverse modulation optical communication has the advantages of a large field of view, low power consumption, and suitability for small platform workers, compared to traditional free-space optical communication [69]. By simplifying the uplink

Table 34.3 LOS link vs. NLOS link [61]

Comparator	LOS link	NLOS link
Communication rate	High	Low
Direction requirements	High	Low
Shadow resistance	Low	Intermediate
System mobility	Low	Intermediate
System complexity	Low	Intermediate
Background light effects	Small	Large
Multipath decay	No	Yes
Channel loss	Low	High

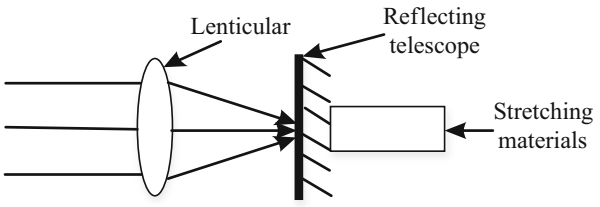


Fig. 34.3 Diagram of a cat's eye structure modulator [67]

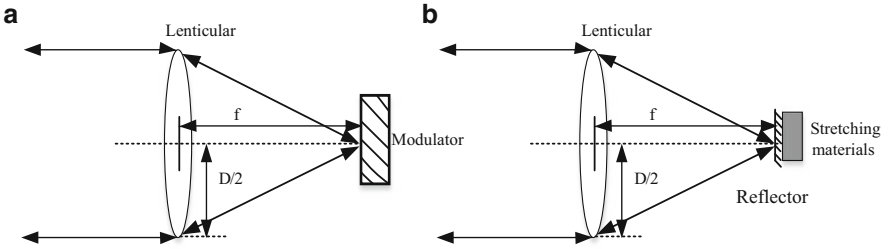


Fig. 34.4 Principle of a cat's eye reverse modulator [68]. (**a**) Non-depolarizing inverse modulator model. (**b**) Off-focus inverse modulator model

using inverse modulation, the complexity of the system structure is greatly reduced, while preserving the inherent advantages of visible light communication.

Equalization Technology

Equalization techniques improve the LED low-pass characteristics and overcome intercode interference by compensating for non-flat channels. Equalizers can be classified into two categories, pre-equalizers (located in the transmitter section) and post-equalizers (located in the receiver section) [70].

Pre-Equalization Technology

Pre-equalization improves the transmission rate and LED response bandwidth by compensating for the signal distortion caused by the device and channel. It can be classified into hardware pre-equalization and software pre-equalization, in terms of the implementation level [71].

Hardware pre-equalization techniques are mainly through the design of analog circuits. Pre-equalization circuit modules are connected in series in the transmitter circuits to counteract the nonlinear fading of the system and achieve the equalization effect [72]. Software equalization, i.e., using digital signal processing methods, automatically adjusts the transmitted signal waveform according to the channel characteristics, to achieve the equalization effect [73].

A finite impulse response (FIR) pre-equalization filter is the current mainstream software pre-equalization scheme; it mainly consists of a window function design method and a frequency sampling design method [71]. The advances in pre-equalization techniques are summarized in Table 34.4.

Post-Equalization Technology

A post-equalization technique is a scheme for equalizing the receiver side of the system to adapt to the time-varying nature of the multipath channel when the transmission environment is constantly changing.

Current analog post-equalization mostly uses passive equalizers or a combination of passive and active equalizers to compensate for the signal [70]. Minh et al. [74] presented the design of a first-order analog post-equalizer. It extended the bandwidth of the system to 50 MHz after equalization, and the communication rate reached 100 Mbit/s. A few years later, Chen et al. [75] proposed new post-equalization circuits, which combined two passive equalizers with an active equalizer and a blue-light filtering technique to obtain a 151-MHz system bandwidth and a 340-Mbit/s transmission rate.

Zhang et al. [76] used a circuit consisting of an op-amp and an R-C network as a post-equalizer to extend the system bandwidth to 220 MHz, which reached a maximum of 553 Mbit/s in OOK transmissions. Hao et al. [77] combined a pre-equalizer with a resonance effect and a third-order active high-pass filter circuit and compensated the system modulation bandwidth up to 42 MHz. Zhao et al. [78] designed a dual cascaded π-type equalizer as a passive post-equalization method. The flatness of the 50-MHz compensated signal modulation bandwidth was 3 dB.

Effective measures have been proposed to solve the problem of distortion caused by the multipath effect of channel software post-equalization, such as feedback-based adaptive equalization mechanisms. Mechanism-based adaptive equalization

Table 34.4 Comparison of pre-equalization technology [72, 73]

Pre-equalization type	Equalization technology	LED morphology	Wideband /MHz	Transmission rate /(Mbit/s)
Hardware	Series resonant circuit	Phosphor	2.5–45	80
Hardware	Transistor T circuit	RGB	6.2–91	477
Hardware	Passive overcorrection circuit	Phosphor	12–77.6	200
Hardware	NPN transistor circuit	Phosphor	3–175	460
Hardware	Improved bias-T circuit	Phosphor	1.0–30	190
Hardware	Single-cascade T amplitude equalizer	RGB	75–125	750
Hardware	Dual cascade equalizer	Phosphor	17–366	1600
Software	Power exponential equalizer	Phosphor	500	2080

and linear filtering–based linear equalization schemes include filter structures, such as linear lateral equalizers and judgment feedback equalizers [73].

Pre- and post-equalization techniques can be combined in visible light communication. Pre-equalization can be used to extend the LED modulation bandwidth at the transmitter side and post-equalization at the receiver side to solve the channel-generated signal distortion problem, thus significantly improving the transmission rate of the system [73].

Information Modulation Technology

Modulation is a key technology to improve communication transmission efficiency. The modulation methods currently applied to VLC systems include traditional modulation methods, such as switch keying and pulse position; derived modulation methods, such as differential pulse position modulation and double-ended pulse-spacing modulation, as well as color shift keying and orthogonal frequency-division multiplexing, etc., which are briefly described below [79].

Digital Pulse Modulation

Digital pulse modulation can achieve high speeds and efficient data transmission with many advantages. Ke et al. [80] of Xi'an University of Science and Technology conducted a detailed analysis and compared the characteristics of various pulse modulation methods in optical communications, including average symbol length, bandwidth requirements, bandwidth utilization, and average transmit power.

To meet the demand for high bandwidth, high power, and low BERs in visible light communication systems, new modulation schemes are constantly being proposed on the basis of mature applied modulation techniques. In practical experiments, the appropriate modulation method can be selected according to the need.

Color Shift Keying

Compared with white light monochromatic VLC systems, multicolor VLC systems can both increase the transmission capacity and modulate the chromaticity of the light source [81]. The visible light communication standard proposes color shift keying (CSK) modulation for different color LEDs, which is designed to solve the lighting compatibility problem in multicolor VLC systems [82].

Color shift keying uses RGB trichromatic light-emitting diodes to represent different transmitted data by different colors of visible light. If we use the binary numbers 00 for red, 01 for blue, and 10 for green, the receiver module will recognize and then demodulate the different colored LEDs into pairs of binary codes. This avoids the nonlinearities of monochromatic light sources due to the flickering intensity, which affects the system performance [83]. A schematic of the color shift keying system is shown in Fig. 34.5 [83, 84].

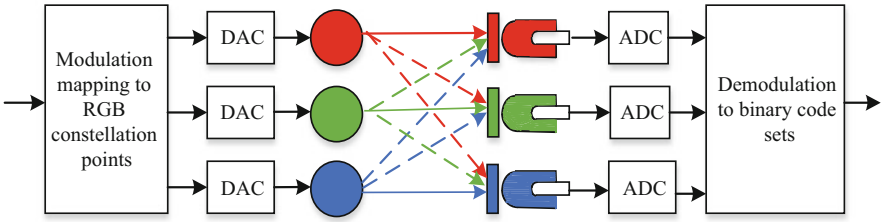


Fig. 34.5 CSK modulation system model

Orthogonal Frequency-Division Multiplexing

The OFDM modulation principle is to modulate a number of signals independent of each other onto carriers of different frequencies and combine them, enabling simultaneous transmissions in the same channel. Because a certain amount of protected band is left between the subchannels, its spectrum utilization is relatively low [85].

OFDM modulation is a type of multicarrier modulation technique, whose basic idea is to divide a given channel in the frequency domain into many orthogonal subchannels; each channel is modulated with one subcarrier to achieve parallel transmissions [85]. As a result, the OFDM spectrum can be overlapped and the original signal can be separated without distortion at the receiver. The use of OFDM is promising because it can combat the intersymbol interference caused by multipath transmissions (Fig. 34.6).

Since 2011, a series of alternative solutions have been proposed by domestic and foreign researchers to address the technical limitations of the high average-to-peak power ratio in OFDM modulation [86].

Bandwidth Extension Technology for White LED Light Sources

The modulation bandwidth of LEDs is defined as the bandwidth between the corresponding frequency point and the reference frequency point, when the output AC optical power drops to half of a certain low-frequency reference value [87]. The lower modulation bandwidth of LEDs in VLC severely limits the development and application of VLC systems.

Auxiliary techniques, such as the optimal design of driver circuits, equalization techniques, OFDM, and channel coding, can improve the system modulation bandwidth; however, the preparation of white LEDs with high luminous efficacy, good color rendering, and large modulation bandwidth can have a multiplier effect. Taking the blue GaN surface coating fluorescence to form white light as an example, the main factors limiting the bandwidth include the following:

The modulation bandwidth of the GaN chip itself.
The radiation lifetime of the phosphor.
The proportion of blue-light flux accounted for in the white spectrum [88].

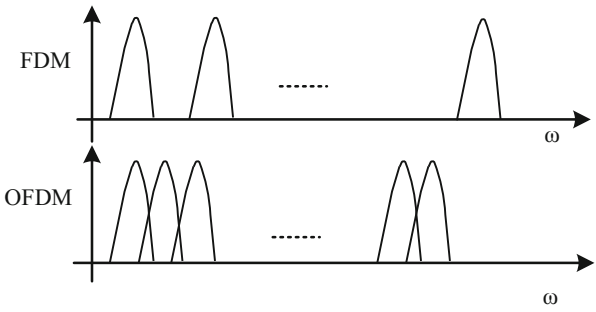


Fig. 34.6 Comparison of the spectrum structure of FDM and OFDM [85]

Pursuing the high-frequency response of white LEDs under high power output and continuously expanding the bandwidth to modulate it up to 30 MHz are the main challenges that must be researched and solved at present for high-speed white LED communication [89].

Optical Camera Communication Technology

Optical camera communication (OCC) is a VLC technique that uses an image sensor (IS) as a receiver to receive optical signals [90]. In existing research, OCC systems mostly use a single large-diameter LED as the transmitting device and a high frame rate camera as the receiving device [91]. The OCC information-sending principle is basically the same as that of VLC, and the high-speed bright and dark flickering signals of LEDs can be used to send the information. The principle of information-receiving principle is to adjust the exposure time of the monocular camera to match the flickering frequency of the LED light source, to capture a clear bright and dark stripe image [92].

The IS used at the receiving end is mainly classified into a charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS). The CCD adopts a global shutter approach, in which each row of its pixels is exposed at the same time. Only one light source state can be captured per exposure, and the communication process encodes only one bit of information, so the global shutter mode data transmission speed is limited.

CMOS uses a roll-up shutter for line-by-line scanning exposure, and the intensity of the light signals presented on different pixel lines is not the same, which means that the information obtained has a multi-bit characteristic. This property can increase the communication rate of visible light imaging [93]. Reference [93] provides a detailed analysis and comparison of modulation methods in OCC systems, which further looks at the development as well as the limitations of OCC. The working process of the two shutters is shown in Fig. 34.7.

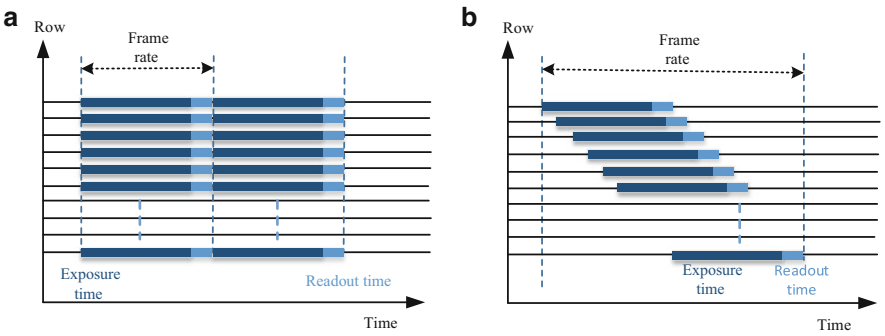


Fig. 34.7 Two types of shutter working processes. (**a**) Global shutter operation process. (**b**) Rolling shutter working process

Indoor Visible Light Communication LED Layouts

In recent years, with the further in-depth study of visible light communication systems, combined with OFDM technology and the development of a number of new schemes, the development of white LED indoor visible light communication has been strongly promoted [38].

Indoor Visible Light System Modeling

An indoor LED visible light communication system uses an LED as a signal transmitter, without affecting the lighting premise by modulating the light intensity of the LED. It relies on high-speed flashing signals that cannot be recognized by the human eye to transmit information [2].

Most current studies on LED light layouts are for 5 m × 5 m × 3 m indoor spaces. This typical indoor visible light system model is shown in Fig. 34.8 to establish an indoor coordinate system, with the LED lights fixed on the ceiling and the receiving end placed on a 0.85-m-high table [90].

Radiant power and luminous intensity are two important parameters of LED light source characteristics. When the LED light source radiation satisfies the Lambert radiation intensity model, its radiation intensity function is [94]

$$R(\varphi) = \frac{m+1}{2\pi} P_S \cos^m \varphi, \varphi \in (-\pi/2, \pi/2), \tag{34.1}$$

where $R(\phi)$ is the radiation intensity when the radiation angle is ϕ, m is the number of radiation modes, and P_S is the LED radiation power. The luminous intensity of the LED light source is

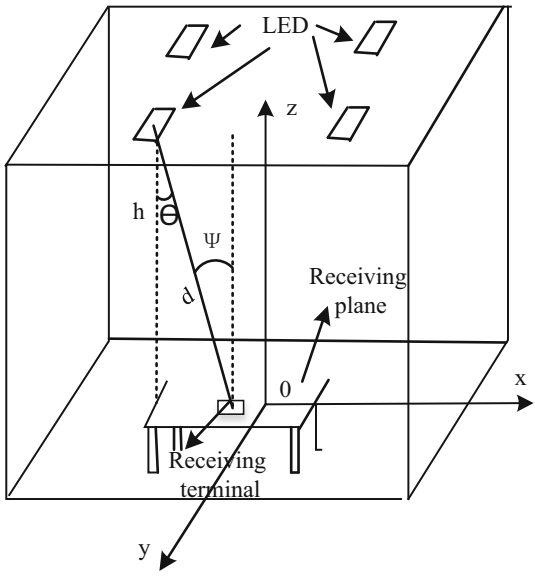


Fig. 34.8 Indoor visible light system model [94]

$$I(\varphi) = \frac{m+1}{2\pi}\varphi_s \cos^m\varphi = I(0)\cos^m\varphi, \tag{34.2}$$

where φ_s is the luminous flux and $I(0)$ is the center light intensity of the LED. The light intensity at a point on the receiving surface is

$$E_{\text{hor}} = I(0)\cos^m\varphi\cos\psi/D^2, \tag{34.3}$$

where ψ is the angle of incidence of the receiving surface and D is the distance between the LED and the receiver.

The light intensity from indoor LEDs to the tabletop is divided into two components: the LOS of the LEDs directed to the receiving end and the NLOS, which contains one or more reflections occurring from the walls or the floor, and then reaches the receiving plane, as shown in Fig. 34.9 [95]. In the non-direct line-of-sight link, the received power distribution is better than in the line-of-sight link; however, the reflections from the walls degrade the SNR [96].

Light Source Layout Design

In a typical indoor LED array layout, the received power distribution does not cover the entire receiving surface; thus, a received power interruption zone can occur, preventing the best performance of the VLC system. To overcome the shadow effect caused by light blindness, a reasonable arrangement of light sources is required.

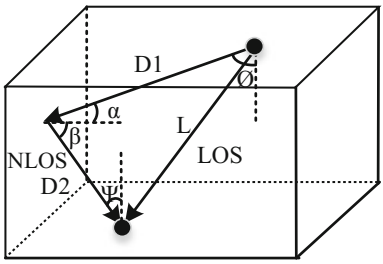


Fig. 34.9 Diagram of LOS and NLOS [95]

Table 34.5 LED light source layout solutions

Year	Layout method	Reference
2007	Four-LED rectangular layout	[97]
2012	Circular angle compensation-type layout	[98]
2012	Single LED in the center + four LEDs in the corners	[99]
2015	16-LED grid layout	[100]
2016	Elliptical layout	[101]
2017	Circular layout	[94]
2017	"4 + 1" rectangular layout program	[102]
2018	Double-nut line layout	[103]
2018	LED array + surrounding ring strip light source layout	[104]
2021	Ring angle compensation-type layout	[105]

However, different LED layout methods can cause uneven illumination and received power [94]. Table 34.5 summarizes several LED layout schemes.

Visible Light Positioning

In recent years, VLC-based positioning technology, i.e., visible light positioning (VLP), has attracted considerable attention, owing to its ability to provide high-precision positioning, based on the existing lighting infrastructure.

Visible Light Positioning Technology

Currently, common localization techniques can be grouped into four main categories: geometric measurement localization (also known as triangulation localization), scene analysis localization (also known as fingerprinting localization), proximity sensing localization, and image sensor localization [106].

In addition, reference [107] proposes a method based on relative illuminance measurements to locate the position of the center of gravity in a circular overlapping area. The illuminance ratio is used to measure the distance between the detector and the projection point of the light source, and the ranging circle equation is constructed

with the projection point as the center of the circle. The center of gravity position of the overlapping region of the triple-ranging circle is derived as an estimation of the detector's position, to achieve accurate localization [107].

The performance of the above visible light positioning methods is compared as shown in Table 34.6, and the most suitable technical solution is selected according to the required accuracy, complexity, synchronization, and other experimental requirements. VLC positioning technologies mentioned in Table 34.6 include angle of arrival (AOA), time of arrival (TOA), time of arrival (TDOA), and received signal strength (RSS).

Positioning Enablement Program

Currently, VLC positioning methods can be divided into two categories based on the type of receiver: photodetector (PD)-based and image sensor (IS)-based. This is because a positioning system with a PD as the receiver will produce large errors, owing to angle measurements, received signal strength measurements, light intensity changes, etc. The image-based method can obtain a more stable and anti-jamming signal, owing to the spatial separation characteristics of the lens in the IS [108]. Moreover, the positioning dimension can be two-dimensional or three-dimensional. The results of the work on visible light positioning in recent years are shown in Table 34.7.

Visible Light Networking

To meet the future needs of multi-user communication and to give full play to the advantages of high-capacity visible light communication, the study of networking technology is timely. Visible light networking is a practical application of visible light communication, and its networking method is one of the main problems it faces.

Networking Overview

Fundamental Principle

Indoor VLC-based networking is mainly divided into homogeneous and heterogeneous networks. Heterogeneous network convergence techniques can ensure that users make full use of network resources and maximize their utility. Therefore, the research on heterogeneous networks is more in-depth [117].

The topology of the networking system is a direct reflection of the connections between the network devices and is also the basis of the system design. Three topologies are proposed in IEEE 802.16.7 for visible optical communication networks, including [118] the peer-to-peer topology, star topology, and broadcast topology, as shown in Fig. 34.10.

Table 34.6 Comparison of visible light positioning techniques [107]

VLC positioning method		Precision	Complexity	Speed	Power consumption	Synchronization	Portability
Triangulation	AOA	Higher	Higher	Slow	High	Yes	Good
	TOA	Higher	Lower	Slow	High	Yes	Good
	TDOA	High	Low	Slow	Low	No	Good
	RSS	Low	High	Slow	High	No	Good
Scenario analysis		Lower	High	Fast	Low	No	Poor
Approximate perception		Depends on network density	Higher	Faster	Low	No	Poor
Illuminance ranging circular overlap area center of gravity positioning method		Higher	Lower	Fast	Low	No	Good

Table 34.7 Visible light positioning technology results

Average positioning accuracy (cm)	Average positioning time (ms)	Number of LEDs	Dimension	Receiver type	Positioning range (L × W × H)m^3	Reference
14.0	/	7	2D	PD	1.5 × 1.5 × 2.5	[109]
0.86	162	4	2D	IS (CMOS)	1.9 × 1.9 × 1.9	[110]
17.52	/	1	3D	IS (CMOS)	3.0 × 3.0 × 2.0	[111]
12.79	/	1	3D	IS (CMOS)	3.0 × 3.0 × 1.5	[111]
2.26	6.3/60	1	2D	IS (CMOS)	1.8 × 1.8 × 3.0	[112]
1	80	2	3D	IS (CMOS)	1.0 × 1.0 × 1.5	[113]
2.5	42	1	3D	IS (CMOS)	8.4 × 4.9 × 2.2	[114]
5.61	64,13	2	3D	IS (CMOS)	1.0 × 1.0 × 3.0	[115]
1.17	/	4	3D	PD	4.0 × 4.0 × 3.0	[116]

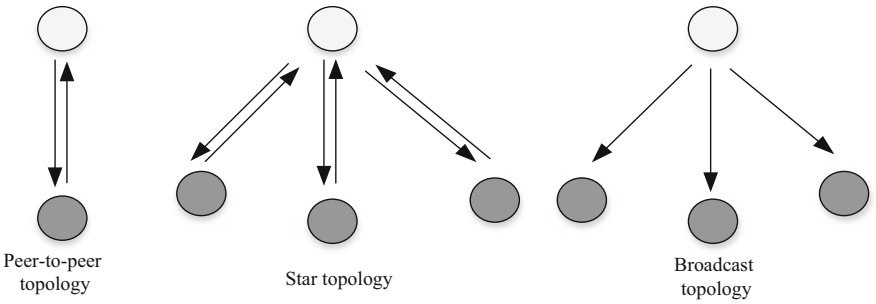


Fig. 34.10 Network topologies [118]

In the peer-to-peer topology, each device can communicate with any other device in its coverage area. A star topology consists of a master node, as well as a number of slave nodes that must pass through the master node in order to communicate. All star networks operate independently of all other star networks currently in operation. Broadcast topology communication is unidirectional: Devices in broadcast mode can transmit signals to other devices without forming a network; therefore, a broadcast network topology does not enable networking.

The star topology places higher functional requirements only on the access point (AP). To ensure low complexity under the premise, as long as rational networking protocol designs are followed, they will be suitable for a wide range of commercially available visible light networking systems.

Multiple Access Technology

Within a visible light communication networking system, multiple access technology enables signals from different users to access the local area network

according to certain rules without interfering with each other. Normally, the AP can only provide service to one terminal at a given moment. Therefore, the network needs to define the MAC layer corresponding to the access protocol to allocate the terminal to use the channel. Table 34.8 shows commonly used visible

Table 34.8 Summary of common multi-access technologies [119]

Category	Name	Principle	Advantages	Disadvantages
Non-contested multiple access	Time-division multiple access	Different terminals send and receive messages in different time slots	High average power efficiency and low realization difficulty	Requires device synchronization; high overhead
	Frequency-division multiple access	Different terminals send and receive messages in different frequency bands	Simultaneous transmission of multiple signals	The more frequency bands are divided, the slower the transmission rate on a single channel
	Code-division multiple access	Different terminal signals are distinguished using orthogonal code-word sequences	High resistance to fading and interference for bandwidth communications	With many users, the code will be too long
	Non-orthogonal multiple access	Different transmit powers are allocated to different terminals	Multi-channel signal transmissions at the same time and same frequency; large transmission rate	Higher complexity
Competition-based multiple access	Carrier listening multiple access	The terminal sends and receives data independently, listens to the channel status through the carrier, and sends data when the channel is idle	Simple to implement, as each site has the same priority and does not require centralized control	When the network load increases, more idle nodes listen to the channel at the same time, causing conflicts and leading to a reduction in network transmission efficiency
	Handshakes	Establishes communication links through request-and-response mode	Timeout retransmission is available; resources are reserved to solve the problem of high conflicts	Control command message exchange takes considerable time and overhead

light communication multiple access schemes and their advantages and disadvantages [119].

Heterogeneous Networking

VLC–Wi-Fi Networking

VLC has a small coverage area and high data rates, while Wi-Fi has the opposite characteristics. In a VLC–Wi-Fi heterogeneous network, the communication areas coexist. Compared with the traditional Wi-Fi-only and VLC-only systems, a VLC–Wi-Fi heterogeneous network provides seamless and reliable coverage and enables high-speed and flexible wireless access [120].

Figure 34.11 shows the four existing VLC–Wi-Fi heterogeneous networking schemes, where MT is a mobile terminal. The link transmitted with VLC has a high-speed transmission rate; when the visible link is interrupted, it can be switched to the same uplink or downlink Wi-Fi link to continue the transmission [121].

In indoor environments, the heterogeneous VLC–Wi-Fi network has the following characteristics:

The demand for downlink data is much greater than the demand for uplink data
The uplink optical link will become visually interfered
The line-of-sight transmission of VLC is susceptible to occlusion interference

Compared with the other three schemes, the scheme shown in Fig. 34.11d is more suitable for an indoor VLC–Wi-Fi heterogeneous network. A downlink with Wi-Fi as a complement to VLC not only provides a high-speed downlink transmission rate but also provides a seamless communication experience. An uplink with Wi-Fi data transmission solves the problems of line-of-sight transmission, as well as the visual interference in VLC uplinks [122].

VLC–PLC Networking

Power line communication (PLC) uses medium- and low-voltage power lines as a communication medium for information transmission. The advantage of PLC is that it can be used anywhere in the home, without the need for additional network wiring. By combining the two technologies, VLC and PLC, a new broadband access system can be effectively used in home networks to promote the development of indoor visible light communication [123]. Combining the two technologies using the newly defined unified MAC layer policy is more difficult. Therefore, different media can be used to maintain independent MACs, exchange information above the MAC layer, and apply unified control and management [124].

A typical fixed-point VLC combination scheme is to place the VLC dedicated transmission device under the VLC light source so that the VLC transmits at a fixed point, as shown in Fig. 34.12a. VLC uplinks and downlinks are used to communicate between the VLC light source and transceiver. The VLC light source is connected to

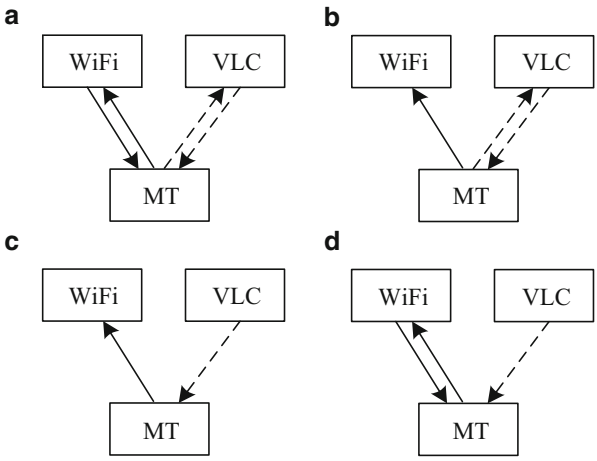


Fig. 34.11 VLC–Wi-Fi heterogeneous network transmission solutions [122]. (**a**) VLC uplink and downlink, Wi-Fi uplink and downlink. (**b**) VLC uplink and downlink, Wi-Fi uplink. (**c**) VLC downlink, Wi-Fi uplink. (**d**) VLC downlink, Wi-Fi uplink and downlink

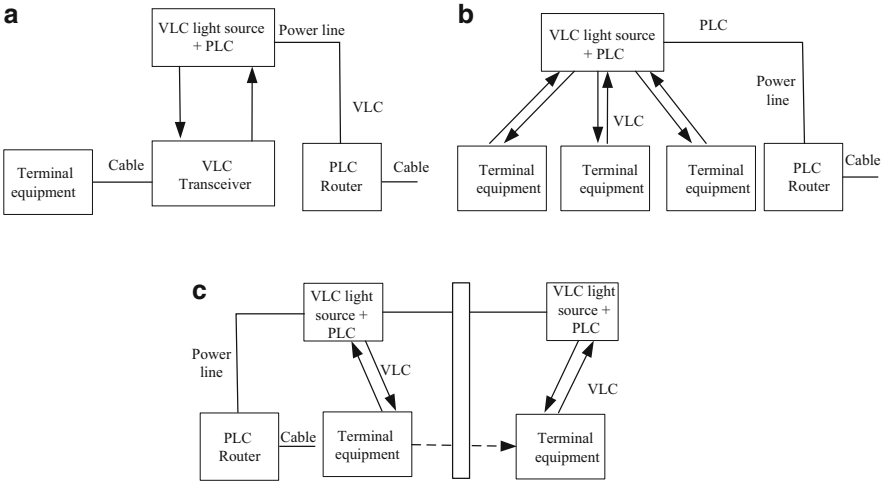


Fig. 34.12 PLC–VLC combination solutions [125]. (**a**) VLC fixed-point transmission. (**b**) Multi-terminal VLC fixed-point transmission (**c**) Multiple VLC light source transceivers

the external network port device by the PLC device, the VLC light source and PLC device are connected by the PLC power line, and both of them use PLC communication [125].

Each Internet access device in the multiple VLC device terminal scheme has a VLC transceiver. The light source is connected to the external network by a power line, and they communicate with each other by PLC, as shown in Fig. 34.12b. The scheme can also be extended to rooms with multiple light sources, where the light sources are networked directly through power lines. The communication

automatically excludes weaker light signals at a distance and communicates using VLC with the nearest light source [125].

The multiple VLC light source transceiver scheme mainly solves the problem of moving a single device from one room to another to ensure uninterrupted communication, as shown in Fig. 34.12c. Because the light source PLC is networked, the new VLC light source remembers the VLC device that is already online through the PLC, so there is no need to re-register the device after it reconnects [125].

Reference [126] provides a thorough review of several technical aspects related to PLC–VLC networking, societal applications, and the most important challenges when designing high-speed integrated PLC–VLC systems [126].

Other Heterogeneous Networking Examples

Intermixing VLC and Ethernet technology uses an optical communication medium conversion module as the core to convert signals between different media of wired network interfaces and wireless free space [127]. Clients can use white light communication to smoothly connect with Ethernet and use custom communication protocols to communicate information between servers and multiple clients and between each client on the integrated network [127].

The fusion of visible light communication and Ethernet is conducive to the construction of a converged communication network of radio, wired optical, optical wireless, and other communication modes; thus, it effectively expands the usability of the network and provides a broad application prospect for indoor short-range communication.

5G networks have always regarded network convergence as an important attribute and have been continuously enhanced and developed. Therefore, the convergence of VLC/5G networks is also an intrinsic need [128]. Reference [129] built an experimental convergence setup, which ultimately led to the successful implementation of an RGB-based VLC system using the new 5G radio standard [129].

Application Scenarios

Visible light communication has a panorama of application development in multiple fields, multiple scenes, and multiple industries. At the present stage, LED visible light communication is mainly used in indoor local area networks (LANs), intelligent traffic systems, and other fields and will develop applications in more directions in the future.

Smart Home

A smart home network is constructed by establishing a smart home integrated software and hardware system based on a visible light communication wireless

LAN. The main research and development focus is on portable visible light communication interface modules, which can be directly embedded in all types of electrical appliances and network access devices, to achieve low-cost, high-rate LAN networking technology that can match the intelligence of commonly used home appliances [130].

Health Care

VLC applications in health care include operating room lighting systems that support real-time data transmission and synchronize all monitoring devices. In the imaging department, ultrasound can be used for real-time calculations for pathology diagnoses using the VLC system. Visible light transmits data using columns of light with a very small angle of dispersion, so it can provide high-speed, high-capacity data passbands to accomplish high-definition and picture-synchronized remote transmission of dynamic medical imaging. Clinics, inpatient hospitals, and waiting areas can be decorated with Internet lights to provide high-speed Internet to attendants. This can also enable doctors to update their medical databases while away from the office [131].

Intelligent Transportation

Visible light communication technology combined with high-speed data image sensor (e.g., high-resolution, high-speed camera) technology enables the system to receive and send timely information about traffic operating conditions. This, in turn, allows traffic personnel to obtain real-time traffic light identification and road condition information, to make appropriate driving adjustments and improve traffic quality and efficiency [132].

Aviation Applications

Because RF communication produces severe electromagnetic interference, it cannot be used in aircraft. Because of the inherent characteristics of LEDs, such as security and the ability to facilitate seamless data transfers, an integrated PLC–VLC system is more suitable than Wi-Fi technology for deployment in aircraft to provide in-flight entertainment, other emergency services, etc [133]

Ship Communication

VLC provides a new type of wireless communication transmission inside a ship by modifying the lighting circuits, solving the problem of serious wireless signal blockage between cabins. By further connecting sensors, VLC can be used to

sound alarms and locate fires by using the signal attenuation caused by flames and smoke obstruction. Using the lighting equipment, the ship's personnel and key equipment can be equipped with optical receivers to localize the auxiliary management and control [134].

Intelligent Industrial Lighting

An intelligent lighting control system using visible light communication technology can not only function as an intelligent lighting management system but also be part of a fast LAN communication network and be combined with the Internet for remote control and access to large amounts of data. The use of an intelligent lighting control system combined with virtual reality technology can provide position information about items in an intelligent warehouse and analyze the detection results in an automated assembly line. Thus, it can lead to the future revolution of visual work interfaces in industrial production to meet the demand for informationization and humanized construction of smart factories [135].

Museum Environment Monitoring

An intelligent museum environmental monitoring system constructed using VLC technology transmits environmental monitoring data through LEDs. It provides the dual functions of lighting and communication, thus effectively reducing the duplication of museum-related monitoring systems and lowering operational costs. It enables the stable monitoring of museum environmental factors, while minimizing damage to museum artifacts [136].

Library Navigation

Indoor navigation in libraries can be realized using the latest visible light communication technology. Through the six steps of inputting the target resources on the mobile phone, getting the flashing information from the camera, decoding the information, obtaining the position of the LED light, locating the starting point and direction of the navigation, and finding the destination information on the path, the library indoor navigation is realized. This provides readers with a fast and efficient method of acquiring books, thus significantly improving the efficiency of book finding and service quality [137].

Underwater Applications

In underwater environments, radio-frequency (RF) transmission rates suffer severe attenuation, which makes RF unable to meet the needs of communication. Owing to

the long delay time of sound waves, the RF bandwidth is limited, and it adversely affects the life of marine organisms. In contrast, underwater visible light communication can overcome attenuation and electromagnetic interference. Research on underwater visible light communication systems has achieved preliminary results [132].

Emergency Applications

In some emergency applications, including defense, power plants, and military, it is vital to maintain great security to protect data from intruders. In addition, the effects of electromagnetic interference (EMI) must be overcome. Therefore, the deployment of integrated PLC–VLC systems has proven to be the best way to ensure secure, as well as EMI-free, communication [133].

Development Trends

To further expand the future application areas of visible light communication, improve its communication performance, and breakthrough application bottlenecks, researchers should continue to explore new technologies and methods. Future research work should focus on the following:

Research and development of new LED devices, to overcome the shortcomings of the light source itself, solve the bandwidth limitation problem, and improve the luminous efficiency. In addition, the cost, control complexity, and other issues should be considered to promote the use of large-scale VLC [138, 139].

Currently, widely used infrared photodetectors are not suitable for VLC; therefore, new high-quality photodetectors suitable for VLC need to be developed in a targeted manner to increase the transmission distance of VLC communications [132].

Transferring data from an illuminated LED light source to a device (i.e., downlinks) is very straightforward. However, transmitting data from devices (mobile terminals) to LEDs (i.e., uplinks) is quite a challenging task. To truly realize visible light dual-full-time communication, uplink implementation techniques must be further explored [132, 133].

The VLC data transmission rate is the biggest challenge for visible LEDs. The high bandwidth utilization and superior anti-interference performance of multiplexing modulation techniques should be explored in depth, while methods, such as efficient filtering techniques and novel filters, should be explored to improve the communication rate [80, 140].

In the analysis and design of VLC systems, external factors that block the channel must be fully considered; thus, it is crucial to establish a visible channel model with practical information. The model can be refined by aiding in the selection of the optimal LED light source layout and luminosity, which reduces the blind zones and minimizes intersymbol interference [132, 141].

Current research on visible light heterogeneous networking is still in its infancy. In the future, it will be necessary to explore the coexistence and compatibility of visible light with terahertz, millimeter wave, microwave wireless communication, and other communication modes [139].

Machine learning algorithms can be applied as advanced signal processing algorithms to VLC to accommodate the complex data processing needs of future systems. Although machine learning-based systems have made some progress, they are still lacking in intelligence. Therefore, more novel machine learning algorithms should be explored in the future to achieve synergy and work together to promote the intelligence of visible light systems [139].

Summary and Outlook

This chapter summarized the current status of research at home and abroad and gave a brief overview of some key technologies involved in the visible light communication system model. For the indoor application of visible light communication, several light source layouts and heterogeneous networking schemes were summarized as follows:

The light-emitting mechanism of LED devices and the basic principles of visible light communication were briefly explained.

Domestic and international research on visible light communication was summarized, and the communication rates involved in the written progress are summarized in a table.

Key visible light technologies were discussed, including link technology classification, existing uplink implementation schemes, inverse modulation techniques, equalization techniques, common information modulation techniques, limitations of light source bandwidth and the extension of its materials, and optical imaging communication schemes.

Based on an indoor visible light system model, a variety of LED layouts and visible light localization schemes were demonstrated.

Indoor visible light heterogeneous networking technologies were further introduced, including four VLC-Wi-Fi heterogeneous networking schemes and three PLC-VLC combination schemes.

The application scope of visible light communication technology covers many fields. Through continuous development, it has achieved very impressive research results, at both theoretical and application levels. It can be foreseen that, after solving the problems of LED devices and dual full-time communication, VLC communication will usher in a period of high-speed development in the future. It will form a complete visible light communication network, accelerate the rapid development of the standardization and industrialization of visible light communication, and optimize and upgrade traditional industries.

References

1. Wang S J, Gao C J, Qi L L, et al. Visible light communication system based on white LED [J]. Laser Journal, 2016, 37(09):69–71.
2. Li Z, Wang F G, Liang J Y, et al. Research progress of LED visible light communication [J]. Journal of Lighting Engineering, 2023, 34(01): 29–40.
3. Yuan C C. Research on some key technologies of visible light communication with white LED [D]. Zhengzhou, Henan: North China University of Water Conservancy and Hydropower, 2016.
4. Zhang Z M. Design of LED visible light communication system based on OOK modulation [D]. Chongqing: Chongqing University of Technology, 2020.
5. Zhang P C. Performance study and theoretical design of visible light LED communication system [D]. Wuhan, Hubei: Huazhong University of Science and Technology, 2011.
6. Zhou B, Xia L. Research on the application of white LED in indoor lighting [J]. Modern Building Electricity, 2013, 4(10):27–32.
7. Liu Z Y. Research on fluorescent composite materials for white light LED [D]. Changchun, Jilin: Changchun University of Science and Technology, 2020.
8. Chi N, Shi M, Hajnal, et al. LiFi:Current status and prospect of visible light communication technology development [J]. Journal of Lighting Engineering, 2019, 30(01):1–9+14.
9. Ren F J, Sun Y K. Research on white LED visible light communication and its key technology [J]. Electronic Quality, 2010, 31(05):20–21+25.
10. Jiang M C. An analysis of visible light communication and its key technologies [J]. Communication World, 2015, 22(16):32–34.
11. Tanaka Y, Haruyama S, Nakagaw M. Wireless optical transmissions with white colored LED for wireless home links [C]//11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. London, UK: IEEE, 2000: 1325–1329.
12. Tanaka Y, Komine T, Haruyama S, et al. Indoor visible communication utilizing plural white LEDs as lighting [C]//12th International Symposium on Wireless Personal Multimedia Communications San Diego, CA, United States: IEEE, 2001: F81–F85.
13. Fan K, Komine T, Tanaka Y, et al. The effect of reflection on indoor visible-light communication system utilizing white LEDs [J]. International. Symposium on Wireless Personal Multimedia Communications, 2002, 2(728): 611–615.
14. Komine T, Nakagawa M. Integrated system of white LED visible-light communication and power-line communication [J]. IEEE Transaction on Consumer Electronics, 2003, 49(1): 71–79.
15. Douseki T. A battery less optical-wireless system with white-LED illumination [C]//15th International Symposium on Personal, Indoor and Mobile Radio Communications. Barcelona, Spain: IEEE, 2004: 2529–2533.
16. Komine T, Jun Hwan Lee, Haruyama S, et al. Adaptive Equalization for Indoor Visible-Light Wireless Communication Systems [C]//2005 Asia-Pacific Conference on Communications. Perth, WA, Australia: IEEE, 2005: 294–298.
17. Sugiyama H, Haruyama S, Nakagawa M. Experimental investigation of modulation method for visible-light communications [J]. IEICE Transactions on Communications, 2006, 12 (89-B): 3393–3400.
18. Afgani M Z, Hass H, Elgalal H, et al. Visible light communication using OFDM [C]//2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities. Barcelona, Spain: IEEE, 2006: 6–134.
19. Sugiyama H, Haruyama S, Nakagawa M. Brightness Control Methods for Illumination and Visible-Light Communication Systems [C]//International Conference on Wireless and Mobile Computing, Networking, and Communications. Yokohama, Japan: IEEE, 2007: 78–84.
20. Komine T, Lee J H, Haruyama S, et al. Adaptive equalization system for visible light wireless communication utilizing multiple white LED lighting equipment [J]. IEEE Transactions on Wireless Communications, 2009, 8(6): 2892–2900.

21. Vucic J, Kottke C, Nerreter S, et al. White Light Wireless Transmission at 200+ Mb/s Net Data Rate by Use of Discrete-Multitone Modulation [J]. Photonics Technology Letters, IEEE, 2009, 21(20): 1511–1513.
22. Vucic J, Kottke C, Nerreter S, et al. 513Mbit/s visible light communications link based on DMT-modulation of a white Led [J]. Journal of Lightwave Technology, 2010, 28(24): 3512–3518.
23. IEEE Standard for Local and Metropolitan Area Networks–Part 15.7: Short-Range Wireless Optical Communication Using Visible Light [S]. New York, USA: IEEE, 2011.
24. Fujimoto N, Mochizuki H. 614Mbit/s OOK-based transmission by the duo binary technique using a single commercially available visible LED for high-speed visible light communications [C]//38th European Conference and Exhibition on Optical Communications. Amsterdam, Netherlands: 2012: 1–3.
25. Bandara K D, Niroopan P, Yeon-Ho C. Improved indoor visible light communication with PAM and RLS decision feedback equalizer [J]. Iete Journal of Research, 2013, 59(6): 672–678.
26. Tsonev D, Chun H, Rajibhandari S, et al. A 3-Gb/s Single-LED OFDM-Based Wireless VLC Link Using a Gallium Nitride μLED [J]. IEEE Photonics Technology Letters, 2014, 26(7): 637–640.
27. Edirisinghe E, Karunarathna P, Dissanyake D, et al. Design and implementation of a bi-directional visible light communication system [C]//2015 IEEE 10th International Conference on Industrial and Information Systems. Sri Lanka: IEEE, 2015: 519–524.
28. Miramirkhani F, Narmanlioglu O, Uysal M, et al. A Mobile Channel Model for VLC and Application to Adaptive System Design [J]. IEEE Communications Letters, 2017, 21(5): 1035–1038.
29. Islim M S, Ferreira R X, He X, et al. Towards 10 Gb/s orthogonal frequency division multiplexing-based visible light communication using a Ga N violet micro-LED [J]. Photonics Research, 2017, 5(2): A35–A43.
30. Ebrahimi F, Ghassemlooy Z, Olyaee S. Investigation of a hybrid OFDM-PWM/PPM visible light communications system [J]. Optics Communications, 2018, 429(50): 65–71.
31. Bian R, Tavakkolnnia I, Hass H. 15.73 Gb/s Visible Light Communication With Off-the-Shelf LEDs [J]. Journal of Lightwave Technology, 2019, 37(10): 2418–2424.
32. Pesek P, Zvanovec S, Chvoika P, et al. Experimental validation of indoor relay-assisted visible light communications for a last-meter access network [J]. Optics Communications, 2019, 451(15): 319–322.
33. Mohapatra S, Satapathy G, Dash P S, et al. Performance Analysis of Visible Light Communication System With Imperfect CSI [J]. IEEE Communications Letters, 2020, 24(02): 2844–2848.
34. Raj R, Jaiswal S, Dixit A. Dimming-Based Modulation Schemes for Visible Light Communication: Spectral Analysis and ISI Mitigation [J]. Fluctuation and Noise Letters, 2021, 20(2): 1777–1798.
35. Aravindan N, Raja S, Selvendran S, et al. 1.3 Gbps OOK modulated phosphorescent white LED using optimized lattice pre-equalization circuit in visible spectrum [J]. Optik, 2021, 242(10): 167–214.
36. Poulose A. Simulation of an Indoor Visible Light Communication System Using Optisystem [J]. Signals, 2022, 3(4): 765–793.
37. Ali A Y, Farhi L, Laila U, et al. Mitigation of Inter-Cell Interference by using MIMO System in Visible Light Communication System [J]. Fluctuation and Noise Letters, 2023, 22(02): 100–107.
38. Hu G Y, Chen C Y, Chen Z.Q. Study on white LED lighting source used as indoor wireless communication [J]. Optical Communication Technology, 2006, 30(07):46–48.
39. Ding D Q, Ke X Z. Research on visible light communication and its key technologies [J]. Semiconductor Optoelectronics, 2006, 31(02):114–117.

40. Yu Z G, Chen C Y, Zhao J, et al. Diversity reception technology in white LED lighting communication system [J]. Optical Communication Technology, 2008, 32(09):52–54.
41. Cao J. N., Liu Z. Y. Angle diversity technology in indoor lighting communication system [J]. Journal of Harbin Engineering University, 2009, 30(11):1295–1300.
42. Ding D Q, Ke X Z. A new indoor VLC channel model based on reflection [J]. Optoelectronics Letters, 2010, 6(04): 295–298.
43. Wang Y, Zhang M, Wang Y, et al. Experimental demonstration of visible light communication based on sub-carrier multiplexing of multiple-input-single-output OFDM [C]//Opto-Electronics and Communications Conference. Busan, Korea (South): IEEE, 2012: 745–746.
44. Wang X, Wang Y, Chi N, et al. Demonstration of 575-Mb/s downlink and 225-Mb/s uplink bi-directional SCM-WDM visible light communication using RGB LED and phosphor-based LED [J]. Optics Express, 2013, 21(1): 1203.
45. Li Y T, Li Z B, Zhang H W, et al. Limiting rate of discrete multi-tone modulated visible optical communication links [J]. Journal of Photonics, 2013, 42(12):1414–1420.
46. Wang C, Zhu Y J, Yang Y H. Research on interleaving division multiple access technology for visible light parallel high-speed communication [J]. Journal of Optics, 2014, 34(04):78–86.
47. Guan W P, Wen S S, Hu H X, et al. Research on visible light communication system based on double modulation technique [J]. Optoelectronics-Laser, 2015, 26(11):2125–2132.
48. Kang Y, Ke X Z. Multi-dimensional coding in visible optical communication [J]. China Laser, 2015, 42(02):138–144.
49. Yang Y, Zeng Z, Feng S, et al. A Simple OFDM Scheme for VLC Systems Based on u-Law Mapping [J]. IEEE Photonics Technology Letters, 2016, 28(6): 641–644.
50. Wang J Y, Li Q L, Zhu J X, et al. Impact of receiver's tilted angle on channel capacity in VLCs [J]. Electronics Letters, 2017, 53(6): 421–423.
51. Ke X Z, Zhang H. Indoor visible uplink communication system based on inverse modulation [J]. Advances in Lasers and Optoelectronics, 2018, 55(02):115–122.
52. Wei L Y, Hsu C W, Chow C W, et al. 40-Gbit/s visible light communication using polarization-multiplexed R/G/B laser diodes with 2-m free-space transmission [C]//Optical Fiber Communication Conference. San Diego, United States: The Optical Society, 2019: 3–7.
53. Wu X, Huang Z, Ji Y. Deep neural network method for channel estimation in visible light communication [J]. Optics Communications, 2020, 462(2020): 1–6.
54. Yang H J, Yang H, Song J. Engineering application of power line visible light two-way voice communication system [J]. Experimental Technology and Management, 2021, 38(11):1–4.
55. Ke X Z, Qin H H, Yang S J, et al. Nighttime background light noise modelling of visible light communication system for vehicle networking [J]. Journal of Radio Wave Science, 2021, 36(06): 986–990.
56. Li Y B, Wei Z X, Wang Z M, et al. THO-OFDM scheme for visible light communication with noise suppression and dimming control [J]. Optics letters, 2022, 47(19): 5180–5183.
57. Chen L B, Li N, Wei J Y. Optimization algorithm of joint sub-channel and power resource allocation for multi-colour visible light communication based on OFDM [J]. Semiconductor Optoelectronics, 2023, 44(02): 294–298.
58. Yang K. Study on nonlinear distortion suppression in visible optical communication system [J]. Optical Technology, 2023, 49(02):184–190.
59. Guang J G. Research on key technology of LED visible light communication link [D]. Harbin, Heilongjiang: Harbin Institute of Technology, 2016.
60. Zhang L. Research on indoor wireless optical communication duplex system based on LED lighting lamp [D]. Zhenjiang, Jiangsu: Jiangsu University, 2010.
61. Zhu J. Research on uplink and downlink of indoor wireless optical communication based on white LED [D]. Zhenjiang, Jiangsu: Jiangsu University, 2008.
62. Cao X B, Wu F, Li J S. Research and prospect of visible light communication and LED lighting system [J]. China Lighting Electric Appliance, 2016, (06): 25–29.

63. Chen Q R, Zhang Tao, Zheng Weibo, et al. Research status and application prospect of visible light communication based on white LED [J]. Semiconductor Optoelectronics, 2016, 37(04): 455–460.
64. Zheng Z, Liu L, Hu W W. Exploration of uplink scheme applied to visible light communication [J]. ZTE Communication Technology, 2014, 20(06):8–11.
65. Zhao X. Key technology of visible light wireless communication [J]. Electronic Technology and Software Engineering, 2015, 4(04):60–61.
66. Feng Q. Research on uplink and downlink of indoor visible wireless LAN based on white LED [D]. Nanchang, Jiangxi: Nanchang University, 2014.
67. Qiu H, Wang J Y, Xu Z Y, et al. Reverse modulation optical communication technology [J]. Military Communication Technology, 2015, 36(02):29–33.
68. Zhang H. Visible light full duplex communication system design [D]. Xi'an: Xi'an University of Technology, 2018.
69. Qiu H, Wang J Y, Xu Z Y. Research on optical communication technology with full-duplex inverse modulation of single light source [J]. Optoelectronic Laser, 2015, 26(12):2314–2319.
70. Wang D, Che Z Q. Research on hardware equalisation technology in visible optical communication technology [J]. China New Communication, 2018, 20(03):8–10.
71. Chi N, Lu X Y, Wang C, et al. High-speed visible light communication based on LED [J]. China Laser, 2017, 44(03):7–18.
72. Song X Q, Wei Y C, Zhao ZixuZ X, et al. Research progress on hardware pre-equalisation technology for high-speed visible light communication [J]. Advances in Lasers and Optoelectronics, 2017, 54(08): 26–36.
73. Wang Y F, Cao F, Deng Z Y, et al. LED nonlinearity compensation and bandwidth expansion techniques in visible light communication [J]. Optoelectronic Engineering, 2020, 47(03): 25–36.
74. Le Minh H, O'brien D, Faulkner G, et al. 100-Mb/s NRZ visible light communications using a postequalized white Led [J]. IEEE Photonics Technology Letters, 2009, 21(15): 1063–1065.
75. Li H L, Chen X B, Huang B J, et al. High bandwidth visible light communications based on a post-equalization circuit [J]. IEEE Photonics Technology Letters, 2014, 26(2): 119–122.
76. Zhang M L, Zhang Y G, Yuan X G. A Visible Light Communications System with 220MHz Bandwidth Based on Pre-emphasis and Post-Equalization Technologies [C]//2014 Asia Communications and Photonics Conference. Shanghai, China: IEEE, 2014: 1–3.
77. Hao H G, Chen P Wang, Niu Y N, et al. Bandwidth expansion scheme for visible light communication system based on equalisation circuit [J]. Optical Communication Technology, 2016, 40(11):48–50.
78. Cui Z, Chen Q R, Zhang T, et al. Design of post-equalisation circuits for visible optical communication systems [J]. Optical Communication Technology, 2017, 41(09):28–30.
79. Song X Q, Jia S J, Zhao Z X, et al. Double-width differential pulse position modulation based on visible light communication [J]. Journal of Armoured Corps Engineering College, 2016, 30(03): 79–82.
80. Ke X Z, Liang J Y, Xu D S, et al. Research progress of pulse position modulation technique for wireless optical communication class [J]. Optoelectronic Engineering, 2022, 49(03):3–21.
81. Chen H, Lai X Z, Chen P, et al. Quadrichromatic LED based mobile phone camera visible light communication [J]. Optics Express, 2018, 26(13): 17132–17144.
82. Huang B Y, Gao Q, Xu Z Y. Joint modulation technique for visible light communication based on multicolour LEDs [J]. ZTE Communication Technology, 2014, 20(06):25–28.
83. Song R, Deng C L, Liu H L. Research on the design method of colour shift keying constellation for visible light communication [J]. Optoelectronic Laser, 2021, 32(04):419–424.
84. Shi T F. Research on colour shift keying modulation technology in visible light communication [D]. Xuzhou, Jiangsu: China University of Mining and Technology, 2022.
85. Luo X Q. OFDM technology in wireless communication [J]. Instrumentation user, 2014, 21(05):14–17.

86. Lv Y T. A technical review of modulation modes for visible light communication [J]. China Science and Technology Information, 2017, 29(22):26–27.
87. Yeh C H, Chow C W, Liu Y F, et al. Simple digital FIR equalizer design for improving the phosphor LED modulation bandwidth in visible light communication [J]. Opt Quant Electron, 2013, 45 (8): 901–905.
88. Ruan C. Preparation of quantum dot white LED and its bandwidth characteristics [D]. Changchun, Jilin: Jilin University, 2017.
89. Huang T T. Key technologies for indoor visible light communication with white LED [J]. Information Communication, 2014, 28(07):204–206.
90. Xiao Y, Guan W, Wen S, et al. The optical bar code detection method based on optical camera communication using discrete Fourier transform [J]. IEEE Access, 2020, 8(6): 123238–123252.
91. Sun Y, Shi W, Shao X, et al. Implementation and decoding method of OCC system based on MIMO [J]. Optoelectronics Letters, 2022, 18(1): 18–23.
92. Zhang Y P, Zhu D Y, Ma J, et al. Metro train localisation based on monocular vision and visible light imaging communication [J]. Advances in Lasers and Optoelectronics, 2022, 59(16): 69–78.
93. Hang S Q, Hou C N, Jia J L, et al. Current status and prospect of optical imaging communication research [J]. Journal of Lighting Engineering, 2020, 31(03): 123–133.
94. Wang J A, Che Y, Lv C, et al. LED layout for indoor visible light communication based on energy optimisation [J]. Journal of Optics, 2017, 37(08):64–76.
95. Shen Z M, Lan T, Wang Y, et al. Simulation analysis of indoor visible light communication system based on LED lamp [J]. Infrared and Laser Engineering, 2015, 44(08): 2496–2500.
96. Hang Y, Yue D W. LED layout for indoor visible light communication based on whale optimisation algorithm [J]. Optoelectronics Technology, 2020, 40(03): 201–208.
97. Ding D Q, Ke X Z, Li J X. Light source layout design and simulation study of VLC system [J]. Optoelectronic Engineering, 2007, 34(01):131–134.
98. Wang Z X, Yu C Y, Zhonh W D, et al. Performance of a novel LED lamp arrangement to reduce SNR fluctuation for multi-user visible light communication systems [J]. Optcs. Express, 2012, 20(4): 4564–4573.
99. Azizan L A, Ab-Rahman M, Jumiran K. Analytical approach on SNR performance of visible light communication for modern lighting layout [C]//2012 International Conference on Innovation Management and Technology Research. Malacca, Malaysia: IEEE, 2012: 332–336.
100. Chow C W, Liu Y, Yeh C H, et al. A practical in-home illumination consideration to reduce data rate fluctuation in visible light communication [J]. IEEE Wireless Communication, 2015, 22(2): 17–23.
101. Wang L, Guo M, Tian H. A visible light communication light source layout model and performance analysis [J]. Laser Journal, 2016, 37(03): 92–94.
102. Zhao L, Peng K. Optimisation of indoor visible light communication light source layout based on white light emitting diode [J]. Journal of Optics, 2017, (07): 49–57.
103. Chen S, Ling L Y, Huang J W, et al. Layout and optimisation of LED light source for indoor visible light communication [J]. Optical Communication Technology, 2018, 42(10):56–59.
104. Zhao L, Zhu T, Liu Z G, et al. A ring light source layout model for balancing illumination and communication [J]. Optoelectronic Engineering, 2018, 45(07):4–9.
105. Jianh H M, Zhou B. Research on light source layout of optical communication system with artificial intelligence technology [J]. Laser Journal, 2021, 42(10):122–126.
106. Zhao A R. Research on Indoor Visible Light Positioning Technology Based on RSS Assisted [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2021.
107. Yang S Q, Zhang Q S, Hu Y, et al. Research on visible light indoor positioning method using illuminance ranging circular overlap area centre of gravity estimation [J]. Optical Technology, 2019, 45(05):557–561.
108. Guang W P. Research on high-precision indoor visible light positioning algorithm based on image sensor [D]. Guangzhou: South China University of Technology, 2019.

109. Li Z T, Yang A Y, Lv H C, et al. Fusion of Visible Light Indoor Positioning and Inertial Navigation Based on Particle Filter [J]. IEEE Photonics Journal, 2017, 9(5): 1–13.
110. Guan W, Chen X, Huang M, et al. High-speed robust dynamic positioning and tracking method based on visual visible light communication using optical flow detection and Bayesian forecast [J]. IEEE Photonics Journal, 2018, 10(3): 1–22.
111. Hu P F, Pathak P H, Zhang H L, et al. High Speed LED-to-Camera Communication using Color Shift Keying with Flicker Mitigation [J]. IEEE Transactions on Mobile Computing, 2019, 19(7): 1603–1617.
112. Li H, Huang H, Xu Y, et al. A fast and high-accuracy real-time visible light positioning system based on single LED lamp with a beacon [J]. IEEE Photonics Journal, 2020, 12(6): 1–12.
113. Guan W, S. Chen, S. Wen, Z, et al., High-accuracy robot indoor localization scheme based on robot operating system using visible light positioning [J]. IIEEE Photonics Journal, 2020, 12(2): 1–16.
114. Guan W, Huang L, Wen S, et al. Robot localization and navigation using visible light positioning and SLAM fusion [J]. Journal of Lightwave Technology, 2021, 39(22): 7040–7051.
115. Xu Y Z, Chen Z. Visible light three-dimensional positioning system based on LED and image sensor [J]. Advances in Lasers and Optoelectronics, 2022, 59(12):1–20.
116. Qin L, Wang D X, Wang F Y, et al. An indoor visible light localisation method based on neural network of extreme learning machine [J]. Advances in Lasers and Optoelectronics, 2022, 59(03): 79–85.
117. Guo W, Yu H Y, Liu J H. Key technology of indoor optoelectronic hybrid network networking based on visible light communication [J]. ZTE Technology, 2014, 20(06):2–7.
118. IEEE Standard for Local and metropolitan area networks–Part 15.7: Short-Range Optical Wireless Communications [S]. New York, USA: IEEE, 2019: 21–2.
119. Li W. Research on Visible Optical Communication Network Architecture and Resource Allocation Mechanism [D]. Beijing: Beijing University of Posts and Telecommunications, 2014.
120. Wang F, Wang Z C, Qian C, et al. MDP-based vertical handover scheme for indoor VLC-WiFi systems [C]//2015 Opto-Electronics and Communications Conference. Beijing: IEEE, 2015: 1–3.
121. Yang C, Xu Y H, Xue L B, et al. Indoor heterogeneous wireless access network based on VLC +WiFi collaboration [J]. Optical Communication Technology, 2016, 40(09):41–43.
122. Wu Y, Zhang S. Research on indoor VLC-WiFi heterogeneous wireless access network [J]. Semiconductor Optoelectronics, 2017, 38(06):853–856.
123. Zong J. Integration of visible light communication technology and power line communication technology [J]. China Science and Technology Industry, 2014, 28(10):36–37.
124. Wang Y, Xu Z Z, Wang N. Combination of power line communication and visible light communication and its mobile application [J]. Telecommunications Express, 2019, 56(11): 5–10.
125. Xu J D, Wang L, Yang L W, et al. Research on combined technology of visible light communication and power line communication [R]. Beijing: China Communications Standards Association, 2014.
126. Vappangi S, Mani V V. A survey on the integration of visible light communication with power line communication: Conception, applications and research challenges [J]. Optik, 2022, 266(77): 1–35.
127. Liu Y, Ca X P, Lin L, et al. Research on LED visible light communication network and Ethernet integration technology [J]. Optical Communication Technology, 2019, 43(01):1–4.
128. Sun T, Lu L,Liu C. Network convergence deepening enables 5G full-scene multidimensional services [J]. ZTE Technology, 2020, 26(03): 56–60.
129. Oliverira M A, Lima E S, Cunha M S P, et al. RGB-based VLC system using 5G NR standard [J]. Optics Communications, 2021, 481(15): 1–7.

130. Lian W Y, Liu J, Cui Sg, et al. Design for The Indoor Visible Light Communication Application System Based on LED Visible Ligh [J]. MATEC Web of Conference, 2017, 139(6): 124–127.
131. Muhammad S, Qasid S H, Rehman S, et al. Visible light communication applications in healthcare [J]. Technol Health Care, 2016, 24(1): 135–138.
132. Zhu Z K. Prospects and development challenges of visible light communication applications [J]. Communication Power Technology, 2019, 36(02):217–218.
133. Matheus L E M, Vieira A B, Vieira L F M, et al. Visible light communication: applications, architecture, standardization and research challenges [J]. In IEEE Communications Surveys & Tutorials, 2019, 21(4): 3204–3237.
134. Liu L L, Hu Y Q, Du Y F, et al. Application concept of visible light communication technology in ship communication [J]. Digital Technology and Application, 2022, 40(04): 1–3.
135. Niu Y F, Sun D,D, Gong T R, et al. Research on the application of visible light communication technology in intelligent industrial lighting [J]. Journal of Lighting Engineering, 2017, 28(05): 27–31.
136. Hai O. Application of IoT technology in museum environment monitoring [J]. Science and technology vision, 2018, 8(17):9–11.
137. Niu Y, Liu Q Rang. Library indoor navigation service based on visible light communication [J]. Library Construction, 2015, 38((10):81–83.
138. Oyewobi Ss, Djouani K, Kurien A M. Visible Light Communications for Internet of Things: Prospects and Approaches, Challenges, Solutions and Future Directions [J]. Technologies. 2022, 10(36):1–18.
139. Chi N, Chen H. Progress of cutting-edge research in high-speed visible light communication [J]. Optoelectronic Engineering, 2020, 47(03): 6–17.
140. Song X Q, Zho Z X, Chen K W. Prospects and development challenges of visible light communication applications [J]. Advances in Lasers and Optoelectronics, 2015, 52(08): 36–44.
141. Zhang J W, Jiao S. Research on the application of visible light in communication technology [J]. Light source and lighting, 2020, 43(12):12–14.

Indoor Visible-Light Communication and Its Heterogeneous Fusion Networks

35

Contents

Introduction 1348
Domestic and Foreign Research Progress 1348
Indoor Visible-Light Communication Systems 1352
  Link Modes and Channel Models 1353
  Uplink Communication Links 1356
Visible-Light Communication Heterogeneous-Fusion Networks 1359
  VLC and Power-Line Communication Integration 1359
  Heterogeneous Integration of VLC and WiFi 1363
  VLC and Ethernet System Integration 1366
  VLC and LTE Network Interconnections 1371
  Interconnection of a VLC and Radio-over-Fiber (RoF) System 1377
Summary 1377
References 1378

Abstract

In order to improve the performance of visible light communication system, the heterogeneous integration of visible light communication technology and other communication technologies can make all kinds of technologies complement each other and enhance the communication system. Unified capacity, improve resource utilization.

Keywords

Visible light communication heterogeneous integration · Ethernet · Power line communication

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_35

Introduction

Visible-light communication is a rapidly developing optical wireless communication technology that not only provides room lighting but also provides users with high-speed data-transmission services. Visible-light communication technology has the potential to design stable communication-network structures in terms of indoor communication systems. The heterogeneous integration of visible-light communication and existing networks enables wireless-access technologies to complement each other, thereby enhancing the network-system capacity and improving resource utilization.

This chapter introduces visible-light communication networks, power-line communication networks, wireless network communication, Ethernet, long-term evolution (LTE) networks, and radio-over-fiber system integration networks. Finally, the development of visible-light communication and heterogeneous fusion networks is discussed.

In recent years, visible-light communication (VLC) has attracted widespread attention as a new type of wireless communication technology. It directly loads the signal onto a light and uses the rapid response characteristics of light-emitting diodes (LEDs) to achieve high-speed signal transmissions while lighting a room [1].

VLC technology has the unique advantages of no spectrum authentication, strong resistance to electromagnetic interference, wide transmission bandwidth, environmental protection, and energy saving. In the future, it will be used for indoor navigation and positioning, provide high-speed information services, and provide wireless-access networks in electromagnetic-sensitive environments, such as hospitals, waiting rooms, and cabins [2]. However, currently, VLC wireless communication networks have the disadvantages of small coverage, immature uplink-transmission link technology, can easily be obscured, limited LED modulation bandwidth, and poor diffraction ability [3].

Therefore, researchers have begun to combine VLC with other wireless-communication technologies to build network systems with large capacities, high-speed information transmission, small access delay, high reliability, strong network security, and flexible structures [4]. Considering the characteristics of various wireless-access technologies, a diversified heterogeneous integrated network architecture can be built to complement all types of wireless-network access technologies. Therefore, research on heterogeneous converged network systems will be an important trend promoting the development of heterogeneous visible-light communication networks in the future.

Domestic and Foreign Research Progress

In 1999, Pang [5] of the University of Hong Kong proposed the concept of visible-light communication technology in the study of LED traffic-lighting systems. He noted that voice-broadcast and traffic-lighting monitoring could be achieved through

light-emitting diodes (LEDs). The entire system data-transmission distance reached 20 m.

In 2000, Tanaka et al. [6] combined optical communication with an LED home link, calculated the received power for the receiver, studied the influence of the size of the field of view on the performance, applied the system to a private network, and created a communication model with an indoor transmission rate of 10 Mbit/s.

In 2002, Zhongchuan Laboratory [7] intensely researched the performance, layout, and channel mode of the lighting source in a VLC system, and analyzed the influence of the signal-to-noise ratio distribution and intersymbol interference in detail, which laid an important foundation for VLC theoretical analysis.

In 2003, Komine and Nakagawa [8] from Keio University proposed a system fusing power-line communication (PLC) and VLC. They provided a specific scheme for two types of network integration, and analyzed their characteristics through simulation. Then, the channel model and noise of the PLC–VLC fusion system were analyzed.

In 2004, Komine et al. [9] studied the influence of shadow and wall multipath reflections on the system in a visible-light channel transmission. They selected a time-division multiple-access modulation method to obtain a data-transmission rate of 800 Mbit/s and effectively reduced the influence of shadows on the system.

In 2005, Komine et al. [10] analyzed a VLC channel from the perspective of line-of-sight (LOS) and non-line-of-sight (NLOS) models and reflection. The performance of the VLC system under on–off keying (OOK), two pulse-position modulation (2-PPM), 4-PPM, and 8-PPM was further studied.

In 2006, Ding et al. [11] of Xi'an University of Technology elaborated on the structure and key technologies of an indoor VLC model, discussed the key issues of light-source layout, receiver-field angle, and intersymbol interference in VLC systems, and provided the corresponding solutions.

In 2007, Ding et al. [12] studied the location deployment of an indoor VLC lighting source, conducted a detailed theoretical analysis of the receiving and light-source power at the receiving end, and simulated the optimal LED-layout scheme. Based on the above scheme, a design scheme for an indoor wireless-communication system using VLC technology was presented, and the key components were analyzed [13].

In 2008, Yun et al. [14] studied VLC modulation and demodulation methods to weaken the influence of intersymbol interference at high data-transmission speeds and proposed a modulation and demodulation scheme based on orthogonal frequency-division multiplexing (OFDM). The simulation showed that the scheme could be used for indoor lighting and high-speed data transmissions.

In 2008, the European Union established the Home Gigabit Access Network (OMEGA) project, and one of its important research areas was VLC technology [15]. Grubor [16] proposed placing a blue filter in front of a white LED, which expanded the modulation bandwidth of high-power white LEDs from 2 MHz to 20 MHz, effectively improving the transmission rate.

In 2008, Chen [17] designed an indoor VLC system-using array LEDs, diversity-reception technology, and intensity-modulation/direct-detection (IM/DD) methods.

The simulation showed that diversity-reception technology was conducive to improving the signal-to-noise ratio of the system. It effectively reduced the impact of indoor-object occlusion and personnel and receiver movement on the communication link.

In 2009, Song et al. [18] used the Monte Carlo ray-tracing algorithm to calculate the impulse response of an indoor visible-light channel. An adaptive OFDM visible-light communication system was proposed. The simulation results showed that the scheme significantly improved the utilization of the frequency-band resources and the overall performance of the system.

In 2009, Komine et al. proposed an adaptive-equalization filter to overcome the intersymbol interference (ISI) caused by multiple lighting sources in indoor VLC. Moreover, an adaptive-equalization filter designed using the least mean squares (LMS) algorithm was used to equalize the indoor VLC multipath channel, and the channel-estimation algorithm was further studied in a mobile environment. The results showed that the adaptive-equalization system more effectively reduced the influence of illumination shadows on the communication network in a mobile environment.

In 2010, Ding et al. [19] first proposed a new mathematical theoretical basic-research model of the LED luminescence mechanism and estimated the model parameters using the Gauss–Newton iterative nonlinear regression algorithm. The results showed that this model could be widely used.

In 2011, Ding et al. [20] designed a new multi-white LED-array transmitting antenna, based on a mathematical model of white LED luminescence, and successfully applied it to an indoor VLC system. They also analyzed the important effects of the number of arrays and their layout, the size of the array optical axis pole angle and the size of the optical array layer spacing in the multi-array light emitting device on the indoor optical power distribution scheme, and accordingly proposed the optimal scheme and guidelines for the configuration of relevant design parameters. Further, the scheme of controlling the optical power at the transmitter is investigated and the new multi-array antenna design with flat-topped optical power distribution is completed.

In 2011, Little et al. [21] of Boston University used WiFi as the uplink for a VLC system. In this scheme, the downlink transmission rate was 10 Mbit/s, and the uplink transmission rate was 30 Mbit/s.

In 2012, Kong et al. [22] proposed placing a number of optical cones in different directions of a hemisphere as an optical receiving antenna, using the angle-diversity receiving method to solve the intersymbol interference on the receiving antenna of a VLC system. A specific layout was designed and implemented, and the total light-intensity receiving problem in the entire hemisphere system was discussed.

In 2013, Chi et al. [23] proposed an indoor two-way communication system using the quadrature amplitude modulation–orthogonal frequency-division multiplexing (QAM-OFDM) modulation method based on red, green, and blue (RGB) LED technology. It combined pre-equalization technology with post-equalization technology. An uplink-transmission rate of 300 Mbit/s and a downlink-transmission rate of 1.15 Gbit/s were realized in the communication process.

In 2014, Zheng et al. [24] analyzed a VLC system with radio frequency (RF), visible light, and 780–950-nm-wavelength infrared light as the uplink communication link and provided a design scheme with a 1550-nm wavelength laser as the uplink communication link. In an indoor environment, a wavelength-division duplex all-optical communication link was formed with a visible-light downlink. Its uplink-transmission rate reached 1 Gbit/s.

In 2014, Tsonev et al. [25] realized a high-performance VLC system with a data-transmission rate of up to 3 Gbit/s on a single LED. They used a 60-MHz modulation-bandwidth gallium arsenide LED and OFDM modulation technology to link the pre-equalization and adaptive data loading.

In 2015, Kang et al. [26] studied 16-PPM multidimensional coding technology and built an RGB LED visible-light audio and video transmission system. The transmission rate on the 3-m communication link reached 1.5 Mbit/s; this multi-dimensional coding technology also effectively enhanced the spectrum utilization of the system and laid a theoretical foundation for high-speed VLC transmissions.

In 2015, Mossaad [27] designed a scheme to solve the problem of a high peak-to-average ratio in the OFDM modulation mode in VLC; that is, the transmitter divided the high peak-to-average ratio signal into several low peak-to-average ratio signals, and then transmitted them by LED. This reduced the influence of the nonlinearity of the LED lighting source on the high peak-to-average ratio signal.

In 2015, Song et al. [28] proposed a single-frequency network (SFN) system design for video broadcasting based on a PLC and VLC deep fusion. OFDM, pre-equalization, and system-equalization techniques were used in the PLC–VLC system. The following year, they studied the channel problem of the PLC–VLC system [29] and analyzed the channel as a whole.

In 2016, Ke et al. [30] comprehensively reviewed visible-light LED communication technology, the basic principle of LED device luminescence, optimal LED location in an indoor VLC system, the array design of the receiving end, and the three link-configuration methods of the VLC system transceiver. Moreover, the bottleneck in VLC technology research was noted.

In 2017, Zhang et al. [31] studied the application effect of user-detection technology in a multi-access VLC system to address the problem of multiple-access interference. They analyzed and understood the effect of correlation detection and blind multi-user detection on the bit-error rate, and summarized a variety of scenarios suitable for VLC applications.

In 2017, Basnayaka [32] designed an indoor VLC system by combining radio frequency and VLC technology to solve the problem of uneven signal coverage caused by uneven illumination in indoor VLC systems. The simulation results showed that the indoor VLC system could greatly reduce the coverage of blind area.

In 2017, Wang et al. [33] proposed an energy-optimal LED arrangement for indoor VLC systems. The particle-swarm optimization algorithm was used to simulate and analyze the distribution of the received power, energy loss, signal-interruption rate, and signal-to-noise ratio under various LED layouts. The simulation verified that the circular-layout scheme required fewer LED lights, optimized

the signal-to-noise ratio distribution, effectively reduced the intersymbol interference of the system, and reduced the overall energy loss of the system.

In 2018, Ke et al. [34] constructed a bidirectional VLC system with a single LED light source using reverse-modulation technology, and proposed a new scheme for constructing an uplink data-transmission link in a VLC system. The uplink and downlink data links of the control system adopted quaternary frequency-shift keying and differential quadrature phase-shift keying modulation, respectively. In the 3-m data-transmission link, the downlink transmission rate was 5 Mbit/s, and the uplink transmission rate was 2 kbit/s. This was the first report on an uplink in this field.

In 2019, Liu et al. [35] clearly proposed technology integrating VLC and Ethernet, and developed a signal-conversion module between wired and wireless media. This integration method achieved full information coverage by the network system. It had high application value in the artificial-intelligence field and other areas.

In 2020, Xue et al. [36] proposed a high spectral-efficiency dimming control scheme by combining OFDM technology and pulse-code modulation (PCM) technology. This scheme not only supported high-speed VLC but also met the requirements of user illumination-source brightness control.

In 2021, Wang et al. [37] applied four-level pulse-amplitude modulation (PAM-4) to a VLC system. An RGB LED was used at the transmitter, and an equalizer was added before the receiver. When the linear equalizer was used at the receiver, the system rate reached 75 Mbit/s.

In 2021, Zhang [38] investigated VLC networking, coding and decoding. He took the lead in building a real-time VLC-based networking-demonstration system platform.

To realize high-speed and safe VLC based on LEDs, the research topics of domestic and foreign scholars include modulation technology, LED light sources, light-source optimization layouts, and other communication systems. To overcome the influence of indoor multipaths, the system bandwidth should be expanded and the transmission rate increased.

Indoor Visible-Light Communication Systems

Visible-light communication (VLC) systems typically include three parts: visible-light transmitters, visible-light transmission systems, and visible-light receivers [39], as shown in Fig. 35.1. The visible-light emitter is mainly composed of a visible-light modulation-coding module and an emission module. The transmitter first encodes, modulates, and pre-equalizes the original data, and then adds an electrical signal to the LED drive current to make the LED glow. At the receiving end, a photodetector (PD) converts the received visible-light signal into an electrical signal, and then the initial signal is restored using demodulation [40].

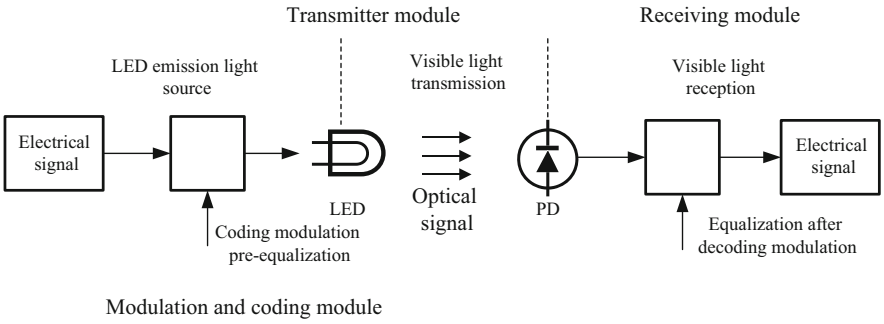


Fig. 35.1 Visible-light communication-system block diagram [40]

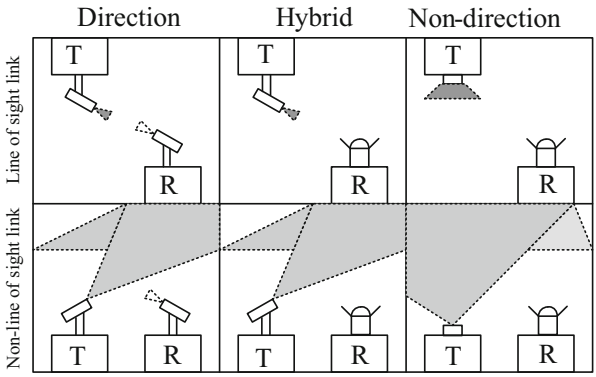


Fig. 35.2 Optical communication-link division standard [42]

Link Modes and Channel Models

Link Modes

In indoor wireless-communication systems, link modes are divided based on two standards [41]. First, the transmission and reception are directional; second, there is no line-of-sight (LOS) transmission. According to the first standard, the transmitting angle of the transmitting end and the receiving-field angle of the receiving end are both very small, and the system link is a directional link. If the sender and receiver do not meet the directional conditions, the system link is a non-directional link; when only the sender or the receiver satisfies the directional condition, it is called a hybrid link [42].

Under the second standard, if the receiver receives both the direct-looking optical signal transmitted directly in the direct-looking space and the reflected optical signal through the indoor emitter, the link is a line-of-sight (LOS) link. If the receiver contains only the reflected-light signal emitted by the indoor emitter, the link is a non-line-of-sight (non-LOS) link [42].

There are six possible link scenarios, as shown in Fig. 35.2.

VLC typically uses the direct-link mode. Following are three variations of the direct link [42].

1. Directional line-of-sight links require the transmitter and receiver to be aligned [42]. Because the optical power consumed by light scattering at the transmitting end is very small, and the angle of view at the receiving end is only focused toward the transmitting direction, the optical receiving power at the receiving end will be large, and the multipath interference will be small under this condition. This method requires a strict alignment between the transmitter and receiver. When the receiving angle is offset or the link is obscured by other obstacles, the communication quality deteriorates significantly. This method is more suitable for static point-to-point communications.
2. Two hybrid line-of-sight chains exist. First, the optical transmitter has a large emission angle and the optical receiver has a small receiving angle. Second, the optical transmitter has a small emission angle and the optical receiver has a large receiving angle. In the first case, an increase in the transmitting angle of the optical transmitter reduces the received power at the receiving end. It also increases the coverage area of the optical communication at the transmitting end to ensure that the user can still communicate with the lamp group when moving the device in a certain coverage area. Because the receiving field of view is very small, the multipath interference generated at the receiving end is significantly reduced. Thus, this scheme is more practical.

   In the second case, the information received by the transmitter is relatively concentrated, and the receiving angle of the receiver is very large. In this case, the receiver undergoes severe noise and intersymbol interference [42].
3. The nondirectional line-of-sight link increases both the transmitting and receiving angles. This method plays a beneficial role in user mobility, so that users can obtain the information sent by the light group over a large range. However, an increase in angle sometimes leads to the enhancement of multipath interference, which seriously affects the link-communication quality [42].

Channel Models

$x(t)$ represents the instantaneous optical signal generated by the LED and $y(t)$ represents the instantaneous current signal of the photodetector. The receiver determines whether data must be demodulated according to the instantaneous current value. Because the multipath signal generated between the transmitting and receiving antennas causes spectral and spatial selective fading, the spatial orientation of the transmitting and receiving antennas determines the amplitude range of the signal.

To obtain accurate results, it is necessary to model the transmission channels. The channel characteristics are represented by a baseband linear system between the input signal $x(t)$, output signal $y(t)$, and impulse response $h(t)$ [30]. The background light, detector, and circuit noise can be regarded as additive white Gaussian noise, $n(t)$, as shown in Fig. 35.3. The channel model is

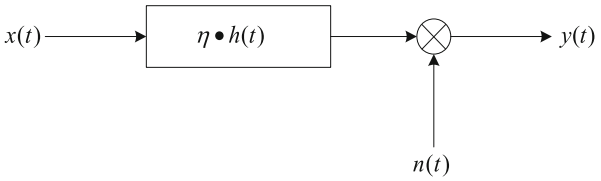


Fig. 35.3 Indoor VLC channel model [30]

$$y(t) = \eta \frac{x(t)}{d^2} \otimes h(t) + n(t), \tag{35.1}$$

where $\otimes$ represents the convolution, η is the conversion efficiency of the photodiode, and d is the distance between the transmitter and receiver. To not lose generality, $1/d^2$ factors are merged into $h(t)$, yielding

$$y(t) = \eta x(t) \otimes h(t) + n(t). \tag{35.2}$$

At present, two main research methods are used for indoor optical wireless channels: the analytical method and the ray-tracing method. Gfeller [43] proposed dividing an indoor surface into several small areas, each of which has Lambert reflection characteristics, and the optical power received by the receiver is the sum of the reflected surfaces. Based on Lambert radiation, Barry [44] proposed a recursive method for solving the transfer function of an indoor channel. The running time of this method increases exponentially with the number of reflections.

The Monte Carlo ray-tracing method is an effective method for studying wireless channels. It studies the channel impulse response, path loss, and time-delay expansion by tracking the path of each ray until the ray reaches the receiving end or exceeds the maximum simulation time. Generally, the probability of a light being received by the detector is very small; therefore, to obtain a reliable simulation result, a large amount of light must be generated, generally between 10^8 and 10^{11} rays. Compared with the analytical method, the ray-tracing method is closer to the actual channel characteristics, and the results are more reliable. However, the reliability is at the expense of a large number of operations.

Rodríguez et al. [45, 46] improved the reflection-propagation model by dividing it into diffuse reflection and specular reflection, and proposed a multiple-reflection model based on the Phong model. An indoor optical wireless channel transmission was modeled using the ray-tracing method, and the error of the impulse response was analyzed.

To obtain better lighting effects, several white LEDs are generally required in a room. Because the LED layout mode is associated with the channel-transmission mode, studying the channel model and its relationship with the location of the lighting source is extremely useful. The LOS link plays a major role in the VLC network, and some necessary measures must be taken to reduce or completely remove the interference of the diffuse optical signal on the system bandwidth. Research on channel models provides an important basis for this.

Uplink Communication Links

In recent years, domestic and foreign research on indoor VLC has mostly focused on key technologies, such as channel modeling, light-source location deployment, modulation methods, coding-method design, and application-system integration development [47]. With the progress of the communication era, the requirements for high-speed and two-way communication in rooms are becoming increasingly urgent.

At present, the technologies for creating communication transmission uplinks mainly include radio frequency, infrared light, lasers, and visible light [48]. Most light-source structures are relatively complex, and damage to the human eye during the movement of the light source is also a problem to be considered.

To address the problems encountered in the design of communication uplinks, researchers have attempted to study reverse-modulation technology. However, current research on reverse-modulation optical communication-system technology is more biased toward basic theory and transmission-link simulations, and reverse-modulation technology is typically used in indoor VLC systems. However, actual cases are rare.

Cat's Eye Reverse Modulator (MRR) is a common reverse modulator that consists of a cat's-eye system and an electro-optic modulator. A modulator is placed at the focal plane of the cat's-eye reverse reflector, and the reflected light intensity is controlled by changing the defocus amount of the transmitting surface to achieve the reverse modulation [49].

In 2017, Ke et al. built an uplink in an indoor VLC system using a defocused cat's-eye MRR and successfully realized full-duplex communication with one-way light sources, simplifying the uplink light source and significantly reducing the complexity of the system implementation. The stability of the network-system transmission was verified experimentally. There was no interference between the uplink and downlink, and the data-return problem of indoor VLC network systems was effectively solved [50].

The link from the reverse end to the active end is the uplink. In [51], the author used the cat's-eye reverse modulation system as the core of the transmission link of the reverse communication system. Figure 35.4 shows the reverse communication-system structure diagram.

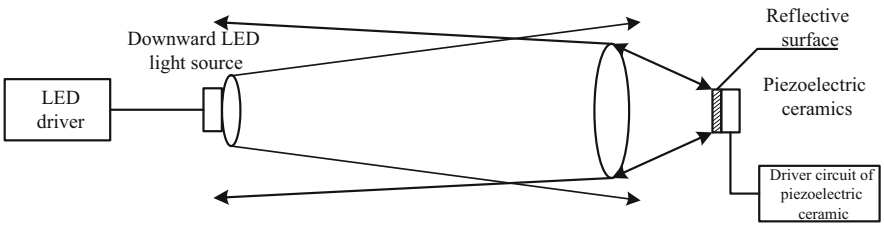


Fig. 35.4 Structure of the reverse communication system [51]

During the system operation, the active end sends the modulated beam directly to the reverse end. The beam-divergence angle of the LED light source is relatively large. The receiving focusing lens located at the reverse end should be discharged side by side with the cat's-eye MRR, and both must be within the effective illumination range of the LED.

The reverse end uses a reverse modulator to modulate the information on the downlink secondary, modulates the signal on the communication uplink to the communication downlink, and then reflects it to the active end. After the receiver at the active end receives the optical signal, the uplink-modulation information is automatically separated and demodulated to realize a duplex-communication link.

The main devices in the uplink hardware include the cat's-eye reverse modulator, drive circuit unit, and signal-receiving and processing circuit unit [52]. The active side must separate the uplink-modulation information from the background light and the information from the downlink high-frequency component. When the modulation rates and modes of the communication uplinks and downlinks are different, the frequencies generated by the two signals passing through the receiver are also different. A band-pass filter is used to separate the uplink signal from the received information, and then the uplink signal is demodulated to restore its uplink baseband-data information. Figure 35.5 shows the uplink-principle block diagram.

In [51], the authors detected the reverse-modulator characteristics using the carrier frequency of the most suitable frequency-shift keying (FSK) modulation method. Based on a preliminary experimental platform, a white LED lamp was used as the lighting source to load the sine-wave signal with different frequencies into the driving circuit of the piezoelectric ceramic sheet. The light intensity of the return signal of the active end was observed using a photoelectric detector when the piezoelectric ceramic sheet had different carrier frequencies. The experimental results are listed in Table 35.1.

From Table 35.1, it can be seen that the light intensity of the return signal is not the same, after a sine-wave signal of different frequencies is input. As the piezoelectric-ceramic plate resonates at a 3-kHz frequency, the piezoelectric-ceramic plate peaks at this frequency. The peak value decreases slowly as the

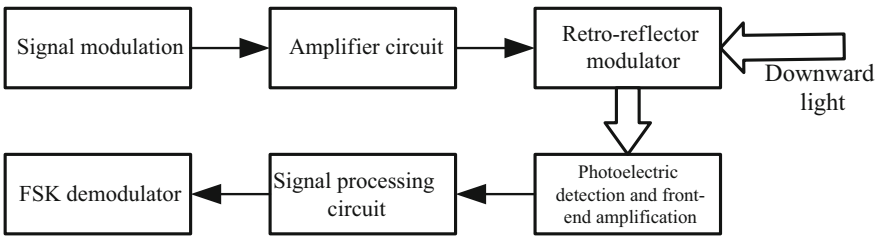


Fig. 35.5 Uplink-principle block diagram [51]

Table 35.1 Peak-to-peak values of return light-intensity signal driven by different frequency signals [51]

Signal frequency (kHz)	1	2	3	12	16	20	Greater frequency
Reflected light-amplitude peak-to-peak value (mV)	10	15	30	10	10	5	Submerged by noise

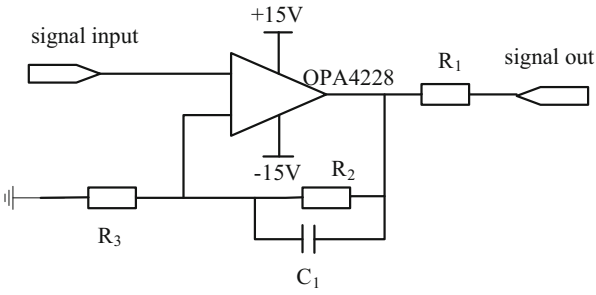


Fig. 35.6 Drive-amplifier circuit module [51]

input frequency of the optical signal gradually deviates from the resonance frequency.

Because the uplink uses a defocused cat's-eye MRR, the defocus amount is proportional to the variation range of the reflected light power. The larger the defocus amount, the greater the change in the return-signal light power, and the better the modulation effect. The defocusing amount depends on the amplitude of the piezoelectric-ceramic sheet, which is proportional to the driving voltage of the modulator [53]. To provide a better modulation effect, the driving circuit must have a wider range of driving-signal outputs. The author chose driving voltages of −15 V–15 V.

The circuit is shown in Fig. 35.6. Because the modulation signal of the digital-to-analog converter (DAC) output is far from sufficient to drive the piezoelectric-ceramic sheet, the signal amplitude must be increased to ensure that the piezoelectric-ceramic sheet operates normally. The OPA4248 amplifier used in the circuit has the advantages of low noise, strong driving force, and large bandwidth; thus, it is used as the driving chip for the piezoelectric-ceramic sheet.

The receiver separates the uplink information from the photodetector after reverse modulation and demodulates it. The reflected light is converted into an electrical signal by a photodetector, but is first amplified to eliminate the influence of background light. Then, a band-pass filter is used to filter out the interference of the communication downlink to extract the uplink information. Finally, the uplink information is amplified and sent to the field-programmable gate array (FPGA) for demodulation [51].

Visible-Light Communication Heterogeneous-Fusion Networks

VLC and Power-Line Communication Integration

VLC uses LEDs as the access points to access indoor broadband data; however, VLC equipment requires power when working, and indoor visible light cannot be directly used as an information source [54]. It must be connected to the backbone network to avoid becoming an information island. Power-line communication (PLC) couples the message to be transmitted to a low-voltage (or medium-voltage) power line through digital modulation, and uses the power line as the transmission medium to complete the transmission.

Because of its good characteristics and lack of additional wiring, it has been widely used in power-line communication systems, the Internet of Things, and smart homes [55]. If PLC technology is integrated into VLC, it will make indoor VLC more convenient for accessing other networks.

Fusion Scheme

Visible-light communication is closely related to power lines and has various applications. Many scholars have closely investigated the combination of PLC and VLC technology and proposed various specific combination schemes [56–58].

VLC is combined with PLC as a wireless technology. It is difficult to define a new unified strategy for the medium-access control (MAC) layer. Therefore, different media are used to exchange information above the MAC layer and establish a unified control-management system [59].

Scheme 1 Fixed-position VLC. This scheme is a classic combination of PLC and VLC. Its feature is that a special VLC transceiver is installed under the VLC light source so that VLC data transmissions can be completed from a fixed position. As shown in Fig. 35.7, Internet-access devices are connected to the VLC transceiver with an Ethernet cable, and the VLC uplink and downlink communicate between the VLC light source and the transceiver. The VLC light source is connected to the external-network port device through the PLC device and connected to the PLC device through the PLC power line; they communicate with each other by PLC. The advantage of this scheme is that the Internet device does not need any modification to connect to the Internet [59].

Scheme 2 Fixed-point transmission VLC. The feature of this scheme is that multiple VLC devices have built-in VLC transceivers, which can simultaneously communicate with VLC light sources. As shown in Fig. 35.8, multiple VLC devices can share a VLC light source in a room, and the VLC light source uses the VLC power line to access the external network through the PLC device, and the VLC light source communicates with the PLC device through the power line. The scheme can also be extended to indoor occasions with multiple VLC lighting sources. These VLC light sources are also connected by power lines, and users can use the nearest VLC light source to access the Internet [59].

Scheme 3 VLC downlink, PLC uplink. To reduce the interference caused by the simultaneous luminescence of multiple VLC devices, each VLC device can be

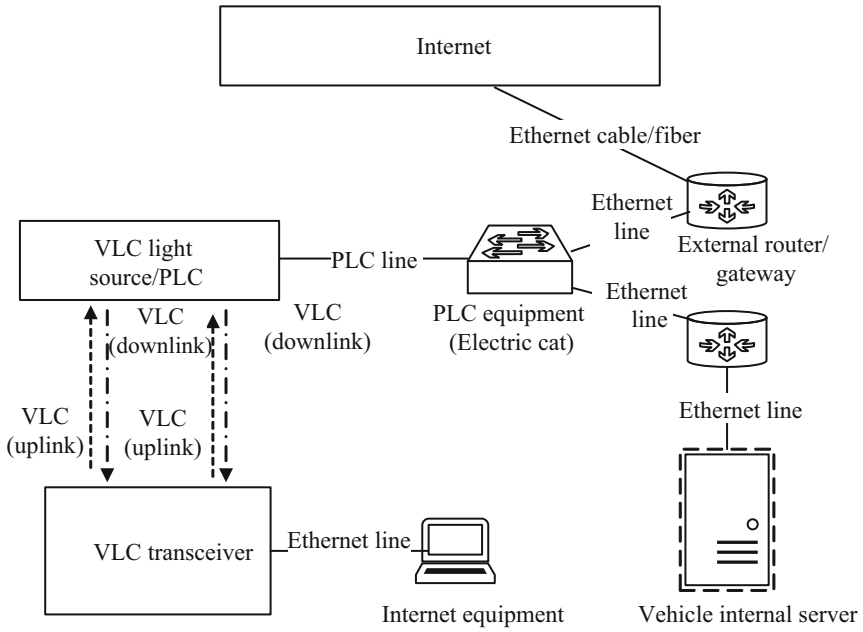


Fig. 35.7 Scheme 1: Fixed-position VLC [59]

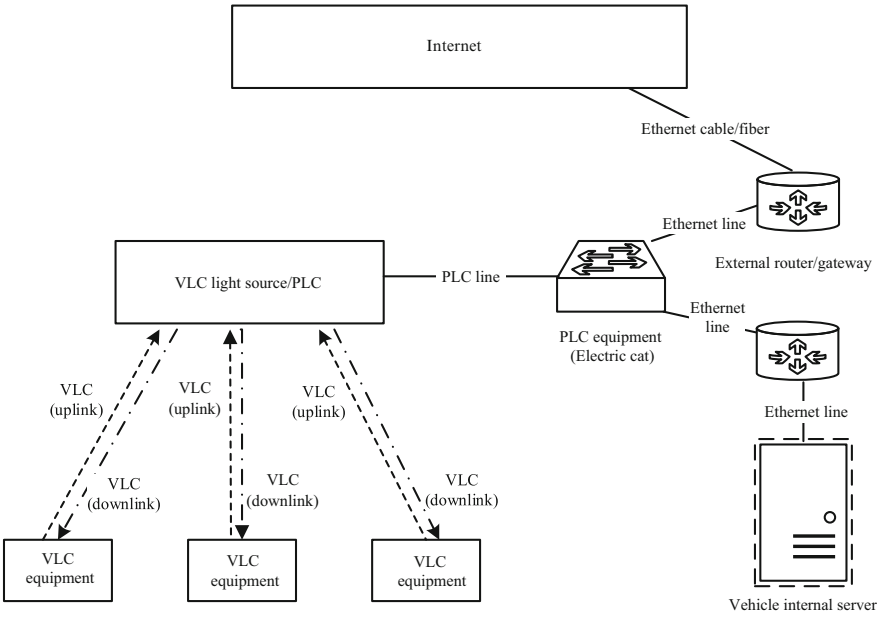


Fig. 35.8 Scheme 2: Fixed-point transmission VLC [59]

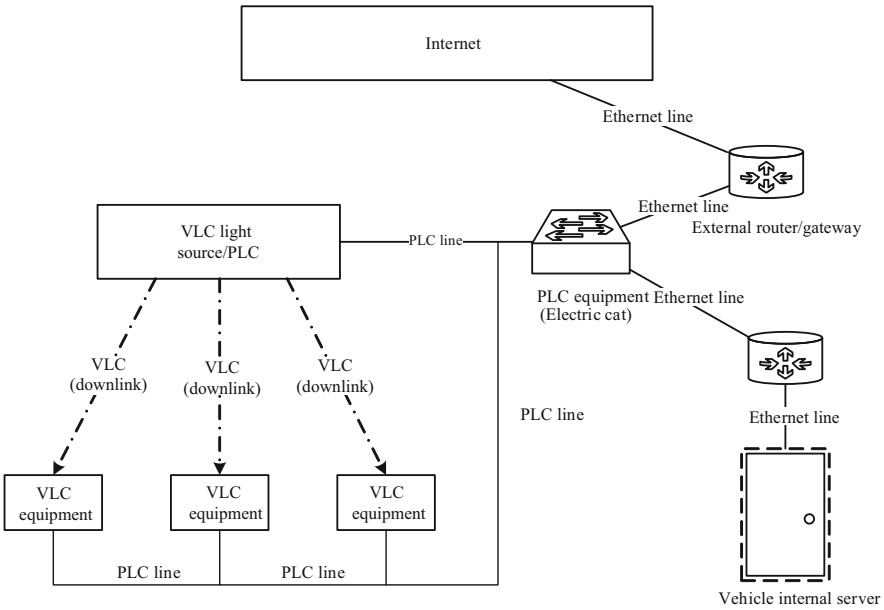


Fig. 35.9 Scheme 3: VLC downlink, PLC uplink [59]

transmitted back by PLC. As shown in Fig. 35.9, the VLC device receives the downlink information from the VLC light source, the power line transmits the uplink information of the VLC light source, and finally connects with the Internet. In addition, the scheme states that to prevent power-line crosstalk between VLC devices, all VLC devices must be connected in parallel on the power line [59].

Fusion-Network Structure

For PLC and VLC fusion systems, the fusion technology is mainly divided into two categories. One demodulates and amplifies the information loaded on the power line, and then loads the LED. The processing method uses a decode-and-forward (DF)/amplify-and-forward (AF) mode [60, 61]. The other loads the LED directly without demodulating the PLC information, which is a deep-fusion method. The principle of the first type of fusion method is shown in Fig. 35.10.

The deep integration of PLC and VLC technology loads power-line information into the LED without modulation and demodulation and uses PLC with visible light as the transmitting point. The PLC signal uses an existing power line to transmit the information and sends a signal to the LED. In addition, no modem is required between the power line and LED [63, 64], as shown in Fig. 35.11.

The above two fusion methods were analyzed, and a general model of the PLC–VLC system was obtained, as shown in Fig. 35.12. PLC/VLC represents the conversion of a PLC signal to a VLC signal.

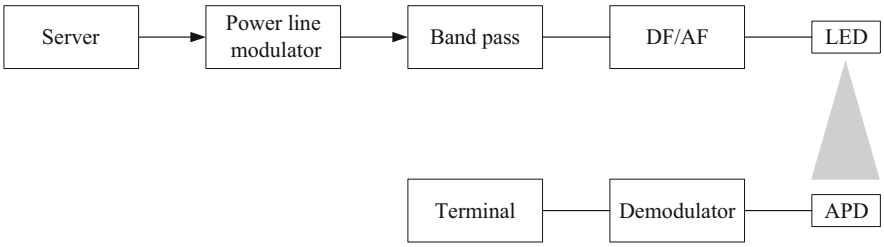


Fig. 35.10 DF/AF fusion mode of a PLC–VLC system [62]

Fig. 35.11 PLC–VLC deep-fusion system [62]

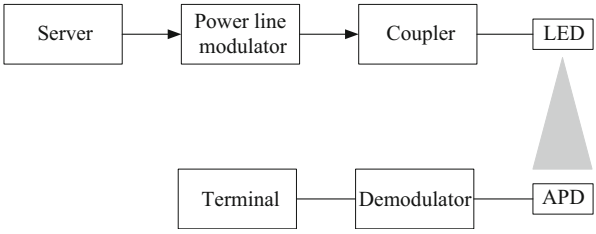
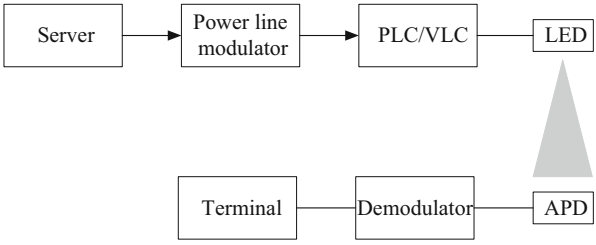


Fig. 35.12 General model of a PLC–VLC system [62]



Channel Model of the Fusion System

The integrated network channels mainly include the PLC channel, LED-lamp channel, VLC channel, and avalanche photodiode (APD) converter. At present, there are two main viewpoints on researching the PLC and VLC fusion-channel characteristics: top-down and bottom-up [65]. In the top-down method, the fusion-channel system is regarded as a black box. Through large-scale test analyses and statistics of certain parameters in the PLC–VLC fusion channel and simulation of the data, a channel model of the PLC–VLC network can be obtained.

The bottom-up method starts with functions that are easier to implement in the network system, gradually analyzes the components that affect the fusion channel from simple to complex, and then uses the two-port network to obtain the channel model of the fusion system, as shown in Fig. 35.13.

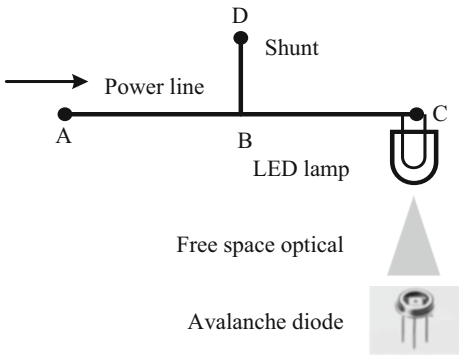


Fig. 35.13 Fusion-system channel model [66]

Heterogeneous Integration of VLC and WiFi

At present, WiFi technology is relatively mature; however, owing to its limited spectrum resources, a single WiFi hotspot cannot meet the requirements of multi-user high-speed Internet access, and the deployment of multiple WiFi hotspots produces co-channel interference [67]. These problems have a serious impact on users' Internet experience. VLC can use light to propagate in a very small space and transmit data to many user devices at a high rate, whereas WiFi can communicate at a relatively low rate in a larger space. Therefore, the combination of VLC and WiFi technologies can complement each other and provide indoor wireless high-speed communication and seamless coverage.

Indoor VLC–WiFi Heterogeneous Networking Scheme

In practical applications, many difficulties exist on the user-demand side when designing VLC transmitter systems, such as the difficulty of integrating the transmitter with an existing intelligent terminal and the limitation of the transmitter's directivity. In addition, because each user may have different transmission requirements for uplinks and downlinks in actual use, other uplink-transmission schemes can also be considered. The WiFi communication rate can reach 300 Mbit/s, and the transmission distance is approximately 100–300 m. The data-transmission distance and rate satisfy the requirements for uplink data transmissions in an indoor-environment network.

A change in the indoor user's location or building pattern may lead to a decrease in VLC signal quality or a disconnection. In [68], the authors used WiFi to assist VLC; that is, the VLC signal was converted to WiFi to transmit data when the VLC signal was interrupted. Four heterogeneous networking transmission schemes are introduced below.

Scheme 1 simultaneously uses visible light and WiFi for uplink and downlink data transmissions [69] and can achieve higher uplink and downlink transmission rates. When visible-light uplink or downlink transmissions are interrupted, the system switches to WiFi to continue the data transmission [68].

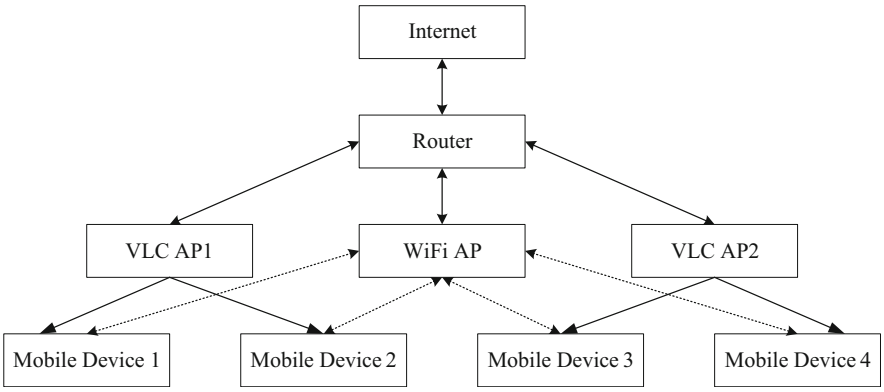


Fig. 35.14 VLC–WiFi topology [68]

In Scheme 2, visible light is used for uplink and downlink transmissions, but WiFi performs uplink transmissions, which can provide high-speed uplink and downlink transmission rates. When the visible-light uplink transmission is interrupted, it can switch to WiFi to continue the transmission [68].

Scheme 3 uses visible light for downlink transmissions and Wi-Fi for uplink transmissions, which provides a high-speed downlink transmission rate [68].

Scheme 4 uses visible light for downlink data transmissions, while WiFi performs uplink and downlink data transmissions, thereby achieving high-speed downlink transmissions. If a visible-light downlink transmission is interrupted, it can be converted to WiFi to continue the data transmission [68]. The topology is shown in Fig. 35.14.

Characteristics of Indoor VLC–WiFi Heterogeneous Networking

Currently, the uplink schemes of indoor VLC–RF heterogeneous networks include visible light, infrared, and WiFi. There are two drawbacks of using visible light as an uplink-transmission signal. First, it is necessary to configure a visible-light-transmitting device at the photodetector (PD) end. Current technology is far from being able to miniaturize visible-light transmitting devices for integrating with other mobile devices. Second, visible light requires a large transmission power and can only be transmitted by an LOS link.

For the terminal PD, it is necessary to solve the problems of power, volume limitation, and location interference to the LOS link [26]. When infrared is used for uplink communication, its half-power angle is very small, usually between 15° and 30°. In addition, the human eye is sensitive to infrared (IR) light. Users usually control their output power in practical applications, and the working range is usually approximately 1 m.

In [70], the author achieved an infrared optical-link transmission rate of 1.152 Mbit/s; however, it could only provide a single service as the communication downlink of a heterogeneous network. This makes it very difficult to offload traffic from visible-light communication.

WiFi has low power consumption, low cost, good compatibility, reliable transmission, and strong networking flexibility. Its radiation distance is approximately 100 m, and the transmission power is 100–200 mW. It can easily achieve stable data transmission in indoor environments. The transmission rate is 300–600 Mbit/s, which meets the requirements of system-uplink transmissions.

In an indoor VLC–WiFi heterogeneous network, the VLC access point provides a very high rate for the downlink transmission of indoor VLC with the advantages of high transmission power and large bandwidth. In addition, the WiFi access point uses its technical characteristics of large coverage area, strong anti-occlusion interference ability, and controllable uplink-transmission ability to realize uplink transmissions in heterogeneous networks [71].

In a WiFi cellular network system, the downlink data traffic far exceeds the uplink data traffic, and the two are asymmetric. In this system, the communication uplinks and downlinks do not interfere with each other during data transmission; that is, they are decoupled. Therefore, the communication uplinks and downlinks of the indoor VLC–WiFi heterogeneous network are also decoupled.

Indoor VLC–WiFi Heterogeneous Network AP Layout

The design scheme of an indoor VLC–WiFi network usually depends on the deployment of the VLC. When planning and deploying indoor lights, the actual lighting needs of the indoor users should be fully considered at the beginning of the design of the whole house. Moreover, the indoor-light intensity must meet the international lighting standard, 300 lx–1500 lx; otherwise, the indoor light will be too strong or too weak to have an impact on human-eye health.

In [72], the author studied the load balancing and power control of VLC–RF heterogeneous networks in a 10 m × 10 m × 3 m room. As shown in Fig. 35.15, the VLC–RF heterogeneous network consists of 16 LED lights and an RF access point. The RF access point (AP) is located in the center of the ceiling, each LED is 2.5 m apart, and the dotted circle indicates the range of visible-light communication LOS when the field of view is 45°. There is no communication-blind area in this regular layout.

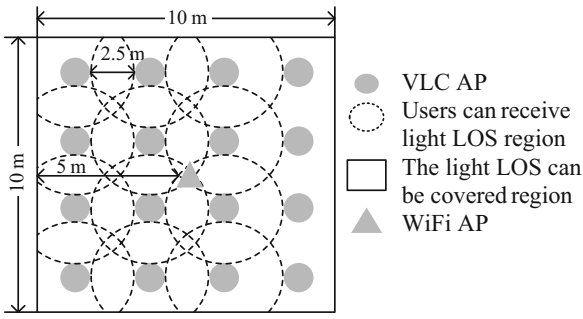


Fig. 35.15 Rectangular uniform AP layout of a VLC–WiFi [73]

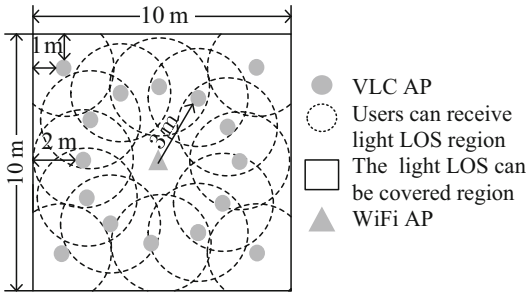


Fig. 35.16 Irregular AP layout in VLC–WiFi [73]

In [73], the author introduced an irregular AP-location deployment scheme to solve the VLC–RF network-access selection problem. As shown in Fig. 35.16, the VLC–RF heterogeneous network consists of 16 LED lights and an RF access point; the RF access point is located at the center of the ceiling. Twelve LED lights are evenly distributed in the central area of the room to form a circle with a 3-m radius, and an LED light source was placed 1 m from the wall at the edge position to supplement the light.

In this scheme, when the LED-emission half-power angle and the user's field angle are both 60°, there are communication blind spots at the center and edge of the room. When the user is in these positions, the communication will be interrupted.

From the above analysis, it can be seen that a reasonable light-source layout will fully meet the lighting needs of the users in the room; however, it is still possible to generate communication blind spots. Therefore, to achieve seamless coverage of indoor communication, an RF access point must be installed in the room.

VLC and Ethernet System Integration

Combining a VLC system with Ethernet technology [74, 75], which has perfect and reliable underlying protocols and standards, will increase the use of indoor-communication access technology in the future. The fusion of VLC and Ethernet technologies accelerates the integration of various communication technologies, such as radio, wired light, and wireless light, and further improves the performance of communication networks.

Two-Way Communication Systems

The client uses white optical communication to transmit messages, as shown in Fig. 35.17. The system communication downlink uses LED visible light, the communication uplink uses a semiconductor laser, and both the communication uplinks and downlinks are equipped with optical transmitters and receivers to facilitate the successful completion of two-way communication [35].

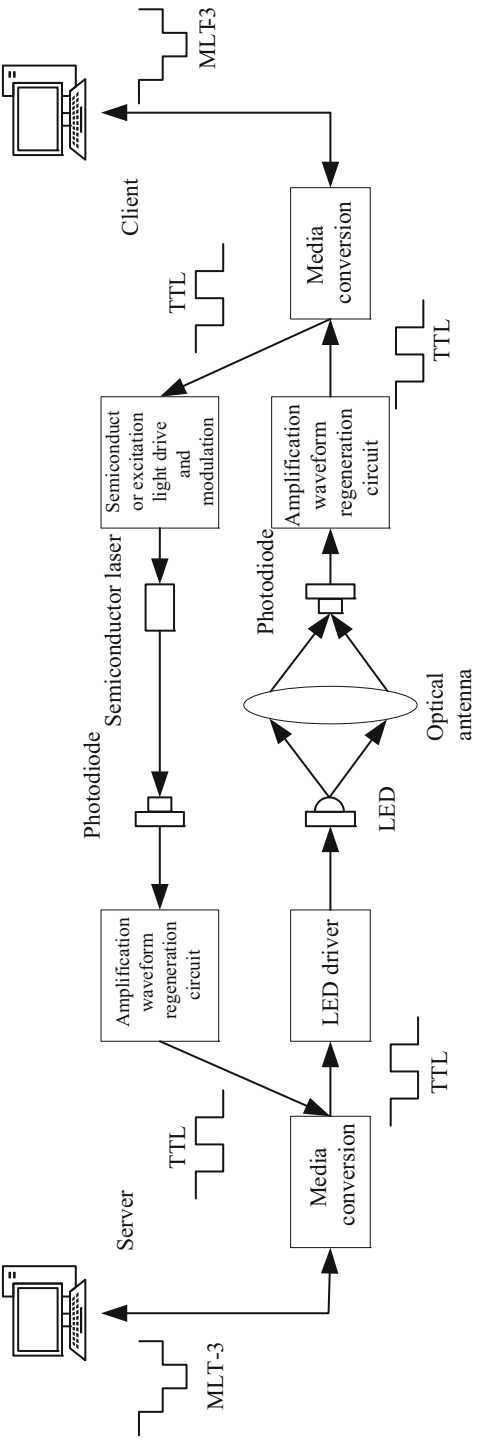


Fig. 35.17 Communication system diagram [76]

The working process of the visible-light downlink is as follows:

1. The server sends the Ethernet multi-level transmit 3 (MLT-3) tri-state bipolar data to the RJ-45 interface-circuit data-processing module and transmits it to the optical-communication medium-conversion module after filtering and coupling.
2. The MLT-3 three-state bipolar level in the network data information is converted to a unipolar TTL level suitable for adjusting the LED through the optical-communication medium-conversion module and transmitted to the LED drive-modulation module.
3. The LED drive-modulation circuit module provides the LED drive current and loads it onto the LED lamp.
4. The optical-communication medium-conversion module converts the electrical signal into network data and transmits it to the client through a twisted-pair cable [76].

Optical Communication Medium-Conversion Module

Network-Interface Circuit

A network-interface circuit generally includes a network control unit and a main processor. The ENC28J60 network control-unit chip, manufactured by Microchip, includes the MAC and PHY layers and supports full-duplex, half-duplex, and other modes. It integrates the physical layer internally and uses a packet-filtering mechanism to constrain the packets sent [77]. Using the industry-standard serial peripheral interface (SPI) [78], the function of the ENC28J60 chip is to pack, unpack, or transform the Ethernet data-frame format.

The network-interface circuit is shown in Fig. 35.18, and a 1:1 pulse transformer is connected to the differential pins (TPIN+/TPIN-) and (TPOUT+/TPOUT-). The

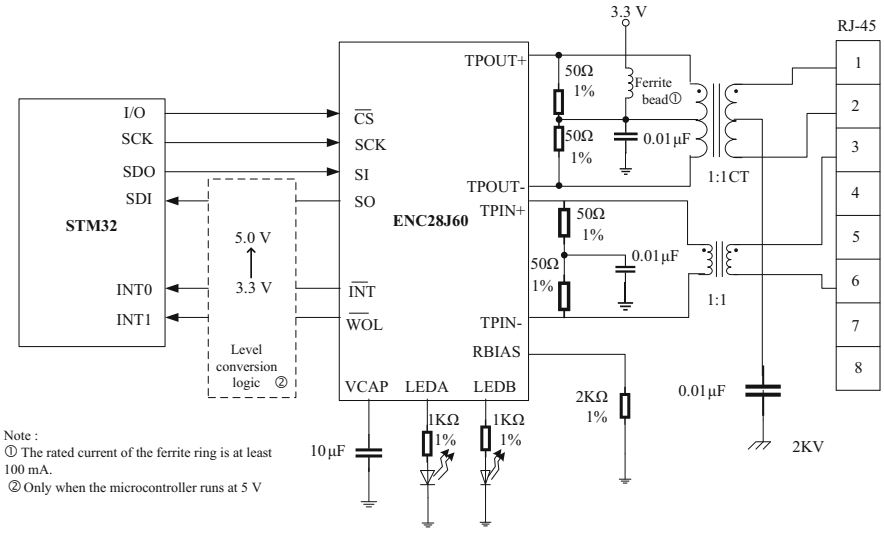


Fig. 35.18 Network-interface circuit [76]

former operates with 10BASE-T without a central tap, and the latter has a central tap. To prevent electrostatic interference, the transformer needs a more than 2-kV voltage isolation, but also must connect to a 50-Ω resistor and 0.01-μF capacitor. The resistance between the RBIAS pin and the ground is 2 kΩ, which meets the requirements of the ENC28J60 ethernet-controller internal analog-circuit design. A 10 μF capacitor is connected between the VCAP pin and the ground to ensure that the regulator works stably.

All the power pins in the interface circuit are connected to a 3.3-V external power supply. In addition, all the reference ground pins are connected to the same external grounding point. The purpose of this design is to drive an external load that the internal 2.5-V regulator of the ENC28J60 cannot directly drive. The LEDA and LEDB pins are connected in series with an LED to display the working state of the circuit.

The optical-communication medium-conversion module uses an STM32F103RBT6 main processor. The main function of the STM32 is to drive the ENC28J60 Ethernet controller to switch the format of the Ethernet data and send and receive Ethernet data information [76].

Protocol Conversion

In an Ethernet system, the signal-coding method is generally the MLT-3 multi-order baseband digital coding method. MLT-3 is a bipolar code with a positive level "+1," zero level "0," and negative level "–1." The information in the VLC is transmitted in space by the change in the TTL level, modulated by the LED, relative to the brightness of the light. The main function of the optical-communication medium-switching module is to switch the polarity level between the two transmission media in the mutual-integration network system and convert the protocol between the communication means [76].

The main functions of the protocol-conversion software given in Reference [76] are hardware and protocol stack initialization, application setting, and packet-format conversion. The working process of the protocol conversion is shown in Fig. 35.19. The Ethernet-frame data packets pass through the data-link, Internet Protocol (IP),

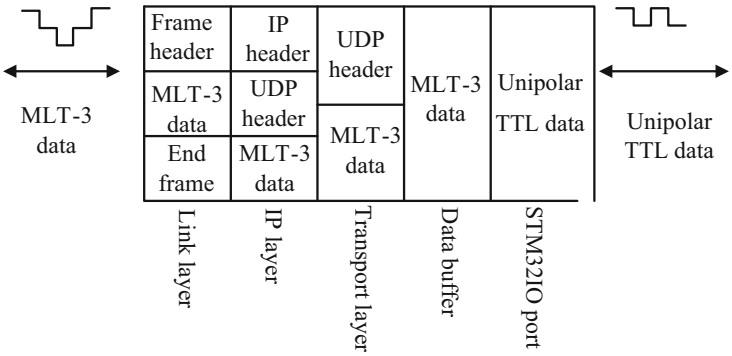


Fig. 35.19 Protocol-conversion workflow [76]

and data-transmission layers, and the frame header, frame tail, IP header, and user datagram protocol (UDP) header are all removed. Subsequently, all the remaining MLT-3 data information is stored in the Ethernet data-cache area, and all of them are converted into unipolar data information through the medium-conversion module [76].

Mutual Fusion of Local Area Networks

The types of network topology are stars, rings, buses, and hybrids [79]. Owing to the common characteristics of VLC and broadcast communication, the star-topology network structure is more suitable for research on interworking network technologies. As shown in Fig. 35.20, this is a common inter-integrated local-area network structure [76]. The network structure includes several clients and a server. Communication between the clients depends on the server, but the server can communicate directly with the clients.

In the Windows system environment, the typical UDP working process is as follows: after starting the server program, the server is loaded into a socket database, and the socket function is established. The local network address must call the Bind function to create a connection. When the server successfully connects to the client, it can call the Sendto and Recvfrom functions for data transmission and reception, respectively. After the transmissions and receptions are completed, both parties call the Close function to close the socket [76].

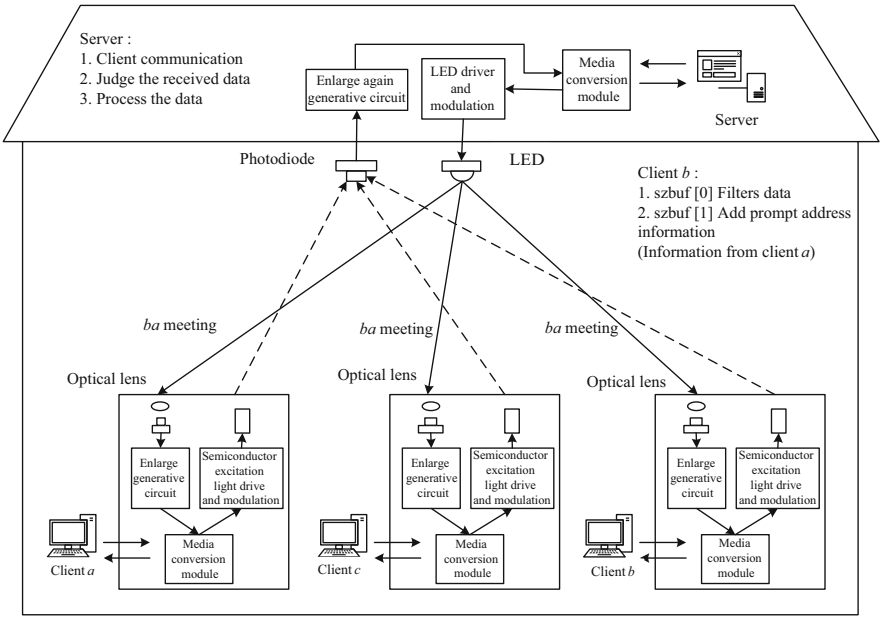


Fig. 35.20 Interconnected network structure [76]

In [35], the author integrates VLC technology and Ethernet technology through an optical-communication medium-conversion module system and makes the two communication technologies complementary to a greater extent; this solves some problems existing in traditional communication technology.

VLC and LTE Network Interconnections

At the Toronto, Canada, Conference in December 2004, the Third Generation Partnership Project (3GPP) officially proposed the long-term evolution (LTE) project [80]. The key technologies of multiple-input multiple-output (MIMO) and OFDM were introduced into the LTE system, which significantly improved the data-transmission rate and system spectrum efficiency. The LTE network supports various bandwidth-allocation schemes, such as 20 MHz, 15 MHz, 10 MHz, 5 MHz, 3 MHz, and 1.4 MHz. It supports the current global mainstream 2G/3G network-frequency bands and new network-frequency bands. The method for allocating spectrum resources is also more flexible.

In terms of network architecture and design, an LTE network has a flatter and more compact network structure. When the number of network nodes is small, the delay is small, and the deployment and maintenance costs of the network are lower. Classic LTE systems include the frequency-division duplexing (FDD)-LTE system and the time-division duplexing (TDD)-LTE system.

At present, many types of wireless communication technologies provide diverse types of services. Bandwidth resources, information-transmission rates, service support, and user mobility have different priorities. Table 35.2 lists the advantages and disadvantages of some representative wireless communication technologies, such as the Global System for Mobile Communications (GSM), wireless local-area networks (WLAN), LTE, and femto-to-femto (F2F).

The future development direction of mobile-communication technology should be to integrate all types of wireless or wired communication technologies so that they can work together to bring seamless coverage and more diversified network services to the majority of users, and fully meet their diverse Internet needs. Based on this concept, some scholars have proposed a heterogeneous network-architecture scheme that combines a VLC network system with other network technologies [81–84]. In [85], the author analyzed the resource-management scheme in a VLC–LTE heterogeneous network based on a virtual cell.

Network Architecture

The positions of the visible-light access points are evenly deployed in the indoor VLC so that the visible-light signal can spread throughout the indoor area. However, when the visible-light transmission channel system is affected by other ambient-light interference, the signal transmission is shielded, network resources are missing, or the access point (AP) is damaged, the VLC system may not work. To ensure that user services are not affected within a certain range, LTE base stations are deployed

Table 35.2 Advantages and disadvantages of different network communication technologies

Communications technology	GSM	WLAN	LTE	F2F
Advantage	The interface standard is universal, and supports encrypted information transmission and multiple services.	The configuration methods are diverse, easy to install and expand, and the cost of network deployment and transformation is low.	The spectrum bandwidth configuration is flexible, cell-edge transmissions are improved, the delay is low, and the rate is fast.	It can make up for the holes in macro cellular networks, improve the signal-coverage quality, enhance the user experience, and reduce the network operation costs.
Disadvantage	The system capacity and data rate are limited, the coding quality is low, and the switching function is poor.	Poor security, information may be stolen, the signal is greatly affected by the environment, and the connection is unstable.	Signaling design is complex, signal penetration is poor, high-frequency signals decay rapidly, and inter-cell interference is difficult to control.	It needs fixed-line broadband resource support, a perfect end-to-end service quality-assurance strategy, and requires users to pay higher fees.

outdoors and all indoor areas can communicate to achieve seamless network switching [85].

In [85], the design scheme of the VLC–LTE heterogeneous network is to evenly arrange several visible-light access points on a 20 m × 20 m × 3 m indoor ceiling, forming a VLC cellular network composed of multiple visible-light access points. The LTE base station is placed outdoors, and the communication range covers the entire indoor area so that indoor users can receive LTE signals.

There is only one access point in the LTE network. According to the actual use of the users, the bandwidth of each AP is set to 20 MHz, and the VLC system has no subchannel-division operation. It uses cooperative transmission and time-division multiplexing technology to provide services to the users. The bandwidth of the LTE network is set to 20 MHz to divide the subchannels. The bandwidth of each subchannel is $B_L = 20/L$ MHz, where L is the number of subchannels. An LTE network uses OFDM technology to provide services to users.

The user can only access one network at a time. $S_N = \{a, b\}$ represents the allowed network set, a represents the VLC network, and b represents the LTE network. $U_k^a = 1$ means that user k accesses network a, and $\delta_k^{am} = 1$ means that user k accesses the mth AP of the VLC network, where $\delta_k^{am} \in \{0, 1\}$. $U_k^b = 1$ indicates that user k accesses network b. To avoid co-channel interference, each LTE network

subchannel is only provided to one user, and user k selects the first subchannel in the LTE network; $\delta_k^{bl} = 1$ and $\delta_k^{bl} \in \{0, 1\}$.

p_k^1 indicates that user k selects the transmission power of subchannel 1 in the LTE network. p_k^a denotes the received signal power of user k in the VLC network, where the received signal power is the power of each AP in the virtual cell received by the user. $B_k^a = 20\text{MHz}$ denotes the bandwidth of user k for each AP. N_0^b is the channel-noise power-spectral density of the LTE network downlink, N_0^a is the channel-noise power-spectral density of the VLC network downlink, and g_k^b represents the channel gain of user k on the LTE network downlink channel 1.

NUM_k^{AP} represents the number of APs in the virtual cell constructed by user k accessing the VLC network and $\text{NUM}_k^{\text{USER}}$ represents the number of users in the virtual cell constructed by user k accessing the VLC network. $\text{NUM}_{\text{VLC}}^{\text{CELL}}$ represents the number of indoor virtual cells.

The signal-transmission rate of user k in the VLC network is

$$C_{\text{cell}}^a = \sum\nolimits_{m=1}^{M=64} \delta_k^{am} B_k^a \log_2\left(1 + \frac{p_k^a}{N_0^a * B_k^a}\right) / \text{NUM}_k^{\text{USER}}. \tag{35.3}$$

The data-transmission rate of a user in the cell cannot be accurately calculated because of joint transmissions and time-division multiplexing in the virtual cell. Therefore, the data-transmission rate of all users in the cell must be calculated and used as the data-transmission rate of the cell.

$$C_{\text{cell}}^a = \sum\nolimits_{\text{num}=1}^{\text{NUM}_{\text{num}}^{\text{USER}}} \sum\nolimits_{m=1}^{M=64} \delta_{\text{num}}^{am} B_{\text{num}}^a \log_2\left(1 + \frac{p_{\text{num}}^a}{N_0^a * B_{\text{num}}^a}\right) \tag{35.4}$$

The signal-transmission rate of user k in the LTE network is

$$C_k^b = \sum\nolimits_{K=1}^{L} \delta_k^{bl} B_L \log_2\left(1 + \frac{p_k^b g_k^b}{N_0^b * B_L}\right). \tag{35.5}$$

The total throughput of the heterogeneous networks is

$$C = \sum\nolimits_{K=1}^{\text{NUM}_{\text{VLC}}^{\text{CELL}}} C_{\text{cell}}^a + \sum\nolimits_{k=1}^{N} U_k^b C_k^b. \tag{35.6}$$

Compared with traditional LTE communication networks, the data-transmission rate of a VLC network is faster, and the bandwidth resources are more abundant. Therefore, when indoors, users may prioritize the VLC network to obtain high-speed and reliable communication services. However, in some special cases, users may also choose to access a traditional LTE network using a vertical handover.

Ideally, users should preferentially access networks with sufficient idle resources, low interference, and low switching costs. However, in practical applications, owing to the limitations of the user's location, the actual occupancy of network resources cannot be fully understood to access the network that meets the requirements.

Therefore, user-terminal access is managed uniformly through the controller of the VLC–LTE heterogeneous network.

Indoor VLC and LTE Network-Interconnection Model

In [86], the authors proposed a method for integrating LTE and VLC networks. As shown in Fig. 35.21, the LTE and VLC network-interconnection system consists of several LTE base stations, VLC access points, service gateways (SGW), and mobility-management entities (MME). The LTE network base stations are connected through the X2 interface. The MME and SGW connect to the entire LTE network base station.

The dual-mode terminal can communicate with the VLC network through its access point and with the LTE network base station through the wireless interface. This system also includes a VLC network scheduler, which is responsible for coordinating the connections between the MME/SGW, LTE network base stations, and VLC access points.

The role of the VLC network scheduler differs on different sides of the system. The LTE base-station side can complete some work closer to the LTE base station; the VLC side is an important device for configuring and managing the entire system. In addition, as an important interface between the LTE network and VLC network equipment, it has a data-information caching function and an exchange-management function. The VLC network scheduler uses the S1 and S2 interfaces to connect the MME and SGW, respectively, and is connected to all the LTE base stations through the X2 interface.

As shown in Fig. 35.22, the terminal is converted from an LTE network to a VLC network. The main steps are as follows:

1. The VLC access-point information is detected when the terminal accesses the LTE base-station network, and the detection report is automatically sent to the LTE base station. The information in the detection report must strictly follow the measurement specifications agreed with the base station, such as the measured

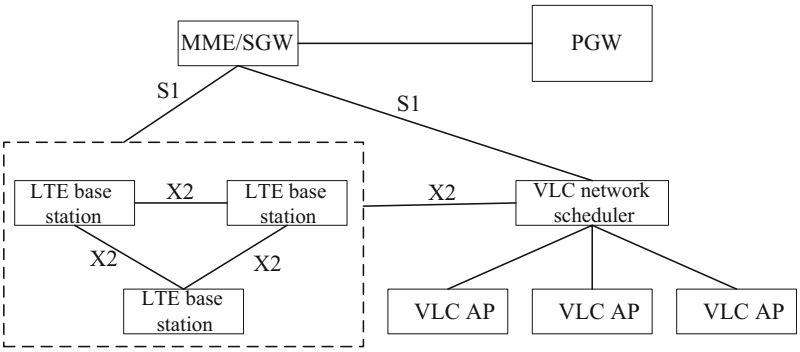


Fig. 35.21 Functional module diagram of the LTE and VLC interconnection system [86]

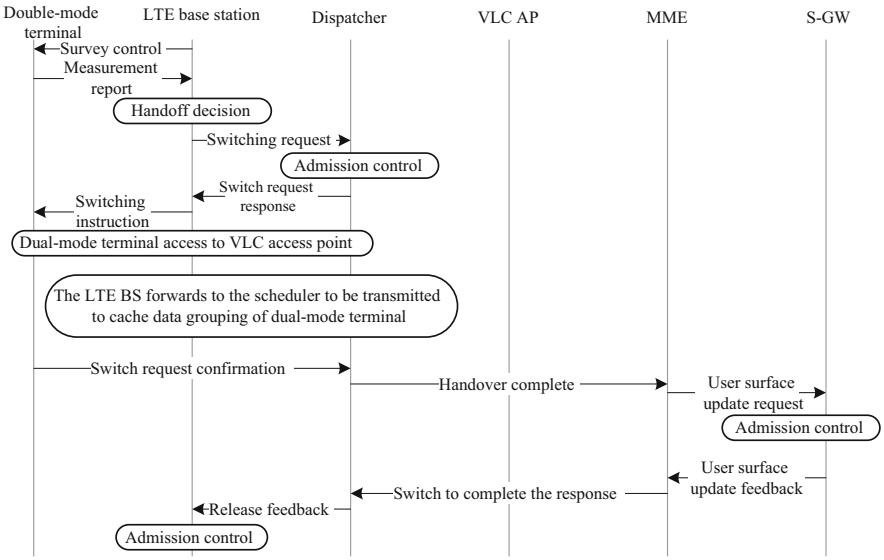


Fig. 35.22 Signaling process for dual-mode terminal switching from LTE to VLC [86]

trigger signal, cycle, and other contents containing cell-identification information and signal-quality requirements.

2. The base station must analyze the data information after receiving the switching signal reported by the user terminal and then determine whether to switch. If the system confirms that it should not switch, it terminates without waiting for the return information; if the base station decides to switch, it sends a switching instruction to the scheduler and enters step 3.
3. When the scheduler receives a handover request, the VLC connects to each VLC access point detected in step 1. After the connection is successful, the scheduler controls the access point and obtains information from it in real time, such as the available resources and load balancing.
4. The scheduler parses the received information related to the access point, determines the admission-control result, and returns the corresponding handover-response information to the base station. The base station sends corresponding instructions to the terminal, and the terminal then determines whether to switch.
5. If the terminal wants to switch, it must access the access point through the handover protocol, which is negotiated by both sides. In addition, the base-station system will be ready, at any time during the entire network-switching process, to save all the data to be sent to the terminal, transmit a copy to the network scheduler, and cache a copy of the local database.
6. After the terminal successfully accesses the VLC, it applies to the scheduler for confirmation and the scheduler sends a corresponding response, according to the current specific state.

7. After the terminal handover, the dispatcher receives the signal at the end of execution and sends it to the MME, which reports it to the SGW, and switches the downlink-transmission signal through the user surface update. When the switching operation is completed, the SGW sends a complete-switching signal to the MME and the MME sends the complete-switching signal to the scheduler.
8. The scheduler requests the base station to release the dual-mode terminal-related information resources after receiving the switching-end signal of step 7. The scheduler then sends the cached data information to the terminal through the access point to ensure a smooth output of the data information.

The process of switching from a VLC network to an LTE network is shown in Fig. 35.23.

The VLC is connected to the LTE through the scheduler. On the LTE base-station side, the scheduler remains completely transparent and has all the characteristics of an ordinary LTE base station. On the VLC network side, all VLC network access points are uniformly managed by the scheduler. The VLC network scheduler also controls the conversion between the LTE and VLC networks.

This scheme has obvious advantages for the LTE system. The conversion process meets the existing LTE requirements and has little impact on the performance of the existing LTE scheme. The VLC network is connected to the LTE network in a tightly coupled manner, which meets the needs of users to a greater extent and reduces the delay caused by system switching, thus ensuring good system performance.

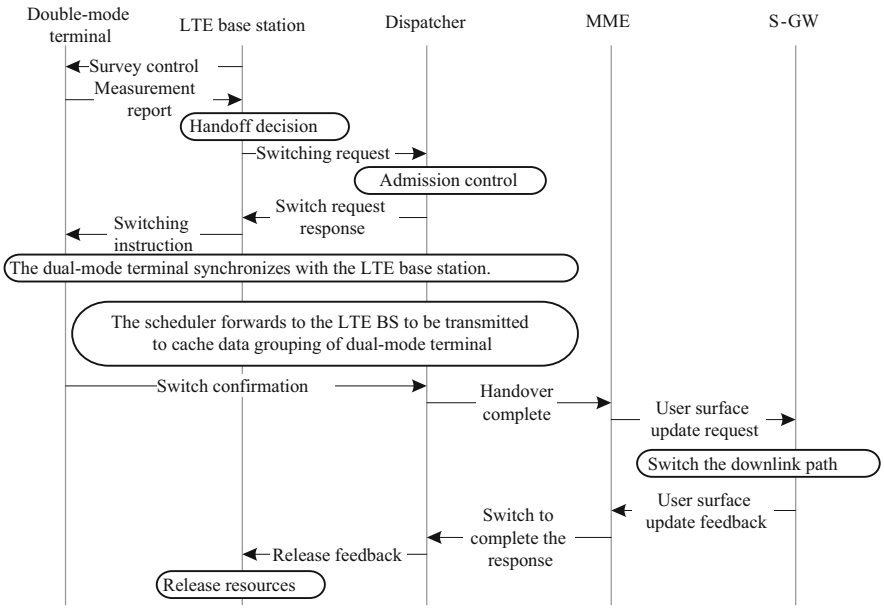


Fig. 35.23 Signaling process for dual-mode terminal switching from VLC to LTE [86]

Interconnection of a VLC and Radio-over-Fiber (RoF) System

The use of white LEDs in lighting systems will become increasingly widespread, and LED-based VLC network technology will also be widely used in the industry. White light sources can be divided into the following two categories:

1. Blue LED light covered with a phosphor layer
2. Mixed red, green, and blue (RGB) light

Comparing the above two types of light sources, the disadvantage of the former is that the bandwidth is small, which is due to the phosphor; for the latter, because the light is filtered, the optical power is reduced. Therefore, the RGB scheme has more advantages for practical applications. However, to achieve high-speed transmissions, modulation techniques, such as discrete multitone (DMT) and OFDM, must be used. In addition, RoF technology can provide a high transmission bandwidth and is more flexible than wired access. If these two technologies are combined, the gains will be different.

When using the mature WiFi IEEE 802.11 g protocol and existing components on the market, RoF and VLC technologies can be integrated. The RoF–VLC system achieves 54 Mbit/s information transmissions. The system is suitable for two-way transmissions, and its performance meets the WiFi standard specifications. Especially for the VLC sublink, the brightness level of the receiver is 15 lx, which is much lower than the standard-lighting brightness level of an indoor office environment.

Summary

With the vigorous development of wireless communication technology, VLC has achieved significant results in basic theory and application fields, which has promoted the rapid development of wireless communication and accelerated the practical application of VLC. The heterogeneous integration of visible-light communication and other communication methods is an indispensable step to meet the needs of future wireless communication. In the future, a major technical challenge for VLC will be successfully integrating existing networks with VLC systems, giving full play to its advantages, and supporting indoor multi-user large-capacity communication. Combined with the evolution and development trends of current visible-light communication systems, the following observations are made:

1. VLC bandwidth resources are very rich; however, the bandwidth of commercial LED devices for modulation is only tens of MHz, which greatly restricts the VLC transmission rate. Therefore, the development of high-speed, high-bandwidth LED devices is a future research and development trend.
2. The visible-light channel-mode selection must be based on the specific light-field distribution and spatial characteristics of the LED component itself. However, in

practical applications, other things must be considered, such as receiver frequency-response characteristics, optical antenna systems, the spatial light-field distribution, atmospheric turbulence, background-light noise, scattering, diffraction, and reflection in the visible-light transmission channel. Integrating this data information to establish a visible-light channel model provides important theoretical guidance for the future development of high-speed VLC technology in China. Moreover, it is of great significance for space and underwater wireless visible-light communication.

3. VLC bidirectional transmission is difficult. Information transmission requires both uplinks and downlinks. It is difficult for a VLC with a commercial LED as the carrier to transmit data upward. It is necessary to increase the LED lamp through a dedicated network card for upward data transmissions, and to add a detector for the downlink-transmission LED to receive the uplink signal. Therefore, using Wiener InGaN/GaN multiple quantum-well LED components to achieve full-duplex communication will become a research hotspot in the future.
4. Currently, VLC heterogeneous networking is still in the primary research stage. The coexistence and compatibility of visible light and terahertz, millimeter wave, and microwave wireless communications will be an inevitable trend in future VLC development.
5. VLC will be a reliable B5G/6G communication method in the future. It can be combined with other communication methods to form a full spectrum network, providing large capacity, high speed, and stable and reliable wireless transmissions for specific scenarios. Examples include intelligent visible-light positioning systems, indoor visible light, and millimeter-wave wireless hybrid networks. To meet the needs of future systems to process complex data, intelligent algorithms based on machine learning should be the focus of attention. They can be used as advanced signal-processing algorithms in VLC systems.
6. In terms of VLC systems, RGB LEDs can provide ultrahigh spectral bandwidth, and wavelength-division multiplexing technology can be used to increase the channel capacity. However, if the color ratio is not well controlled, the human eye will live for a long time in a space that deviates from natural color, which could lead to unforeseen vision problems.

References

1. Luo Hongtu, Chen Changying, Fu Qian, et al. Key Technologies of White LED Indoor Visible Light Communication [J]. Optical Communication Technology, 2011, 35 (2): 56–59.
2. Abdallah K, Sihua Shao, Ammar G, et al. A hybrid RF-VLC system for energy efficient wireless access [J]. IEEE Transactions on Green Communications and Networking, 2018, 2(4):932–944.
3. Liu Jianhui. Research and Implementation of Key Technologies of Indoor Visible Light Communication Network [D]. [Master dissertation]. PLA University of Information Engineering, 2015.

4. Wu Lan. Research on user access and resource allocation algorithm in indoor visible light communication heterogeneous network [D]. [Master dissertation]. Chongqing University of Posts and Telecommunications, 2019.
5. Grantham Kwok-Hung P, Ka-Lim H, Thomas T, et al. Visible light communication for audio systems [J]. IEEE Transactions on Consumer Electronics, 1999, 45(4):1112–1118.
6. Komine T, Lee J H, Haruyama S, et al. Adaptive Equalization System for Visible Light Wireless Communication Utilizing Multiple White LED Lighting Equipment [J]. Wireless Communications, IEEE Transactions on, 2009, 8(6):2892–2900.
7. Komine T, Lee J, Haruyama S, et al. Adaptive equalization for indoor visible-light wireless communication systems [C]. In: Communications, 2005 Asia-Pacific Conference on. 2005, 294–298.
8. Komine T, Nakagawa M. Integrated system of white LED visible-light communication and power-line communication [J]. IEEE Transactions on Consumer Electronics, 2003, 49(1): 71–79.
9. Komine T, Nakagawa M. Fundamental analysis for visible-light communication system using LED lights [J]. Consumer Electronics, IEEE Transactions on, 2004, 50(1):100–107.
10. Komine T, Lee J, Haruyama S, et al. Adaptive equalization for indoor visible light wireless communication systems [C]. In: Communications, 2005 Asia-Pacific Conference on, 2005, 294–298.
11. Ding Deqiang, Ke Xizheng. Research on Visible Light Communication and Its Key Technologies [J]. Semiconductor Optoelectronics, 2006, 27(02): 114–117.
12. Ding Deqiang, Ke Xizheng, Li Jianxun. Light Source Layout Design and Simulation of VLC System [J]. Opto-Electronic Engineering, 2007, 34(1): 131–146.
13. Ding Deqiang, Ke Xizheng. Design and Simulation of Wireless LAN System Based on Visible Light Communication [J]. Journal of Xi'an University of Technology, 2007, 23(1): 29–39.
14. Yun Wenyue. Research on Modulation and Demodulation Technology of Indoor Visible Light Communication System [D]. [Master dissertation]. Changchun University of Technology, 2008.
15. Song Xiaoqing, Zhao Zixu, Chen Kewei, et al. Application Prospect and Development Challenge of Visible Light Communication [J]. Laser & Optoelectronics Progress, 2015, 52(8): 30–38.
16. O'Brien D, Zeng L, Minh L H, et al. Visible light communications [J]. ZTE Technology Journal, 2013, 4(3): 1337–1338.
17. Yu Zhigang, Chen Changying, Zhao Jun, et al. The diversity reception technology in visible-light LED illumination and communication system [J]. Optical Communication Technology, 2008, 32(9): 55–57.
18. Liu Dongyan. Research on indoor visible light communication system based on OFDM [D]. [Master dissertation]. Changchun University of Technology, 2009.
19. Ding Deqiang, Ke Xizheng. Research on Generalized Mathematic Radiation Model for White LED [J]. Acta Optic Sinica, 2010, 30(09): 2536–2540.
20. Ding Deqiang, Ke Xizheng, Li Jianxun. Design of Multi-element Transmitting Antenna for Indoor Visible Light Communication [J]. Journal of Applied Sciences, 2011, 29(03): 238–242.
21. Rahaim M B, Vegni A M, Little T D C. A hybrid radio frequency and broadcast visible light communication system [J]. IEEE Globecom Workshops, 2011, 10(4):792–796.
22. Kong Meimei, Liang Zhongcheng, Zhang Guohu. Optical receiving antenna design based on angle diversity technology [J]. Infrared and Laser Engineering, 2012, 41(03): 750–754.
23. Chi N, Lu X, Huang X, et al. Ultra-high-speed single red-green-blue light-emitting diode-based visible light communication system utilizing advanced modulation formats [J]. Chinese Optics Letters, 2014, 12(1):22–25.
24. Zheng Zhong, Liu Lu, Hu Weiwei. Analysis of Uplink Schemes for Visible-Light Communication [J]. ZTE Technology Journal, 2014, 20(06): 8–11+20.
25. Tsonev D, Chun H, Rajbhandari S, et al. A 3-Gb/s Single-LED OFDM-Based Wireless VLC Link Using a Gallium Nitride [J]. IEEE Photonics Technology Letters, 2014, 26(7):637–640.

26. Kang Ye, Ke Xizheng. Multi-Dimensions Encoding Scheme for Visible Light Communication System [J]. Chinese Journal of Lasers, 2015, 42(02): 138–144.
27. Mossaad M S A, Hranilovic S, Lampe L. Visible Light Communications Using OFDM and Multiple LEDs [J]. IEEE Transactions on Communications, 2015, 63 (11):4304–4313.
28. Song J, Ding W, Yang F, et al. An indoor broadband broadcasting system based on PLC and VLC [J]. IEEE Transactions on Broadcasting, 2015, 61(2):299–308.
29. Ding W, Yang F, Yang H, et al. A hybrid power line and visible light communication system for indoor hospital applications [J]. Computers in Industry, 2015, 68(01):170–178.
30. Xingzheng Ke, Ke Dong. Optical Wireless Communication [M], Beijing Springer Press, 2022: 77–83, 127–140.
31. Zhang Ying, Wang Yansheng. The multiple access and application development of visible light communication [J]. Microcomputer & Its Applications, 2017, 36(18): 18–21.
32. Basnayaka D A, Haas H. Design and Analysis of a Hybrid Radio Frequency and Visible Light Communication System [J]. IEEE Transactions on Communications, 2017, 65(10):4334–4347.
33. Wang Jiaan, Che Ying, Lv Chao, et al. LED Layout for Indoor Visible Light Communication Based on Energy Optimization [J]. Acta Optic Sinica, 2017, 37(08): 64–76.
34. Ke Xizheng, Zhang Huan. Indoor Visible Light Uplink Communication System Based on Inverse Modulation [J]. Laser & Optoelectronics Progress, 2018, 55(02): 115–122.
35. Liu Yang, Cai Xiping, Lin Lang, et al. Research on the fusion technology of LED visible optical communication network with ethernet [J]. Optical Communication Technology, 2019, 43(3): 1–4.
36. Xue Xiaomei, Li Baolong, Li Zhengquan. Research on hybrid visible communication system based on MPPM dimming control [J]. Journal of Optoelectronics·Laser, 2020, 31(06): 648–653.
37. Wang Li, Wang Xuanzheng, Kang Jian, et al. A 75-Mb/s RGB PAM-4 visible light communication transceiver system with pre-and post-equalization [J]. Journal of Lightwave Technology, 2020, 39(5):1381–1390.
38. Chen Q, Han D, Zhang M, et al. Design and Demonstration of a TDD-Based Central-Coordinated Resource-Reserved Multiple Access (CRMA) Scheme for Bidirectional VLC Networking [J]. IEEE Access, 2021, 78(9):7856–7868.
39. Chi Nan, Shi Meng, Ha Yinaer, et al. LiFi: Development Status and Prospects of Visible Light Communication Technology [J]. China Illuminating Engineering Journal, 2019, 30(01): 1–10.
40. Lin Zhenyu. Research on Multi-User Downlink Optimization Method for Indoor VLC-WiFi Heterogeneous Networks [D]. [Master dissertation]. Chongqing University of Posts and Telecommunications, 2020.
41. Kahn J, Barry J. Wireless Infrared Communications [J]. Proceedings of IEEE, 1994, 85(2): 265–298.
42. Huang Xin. Research on Networking Technology of Indoor Visible Light Communication System [D]. [Master dissertation]. Southeast University, 2015.
43. Gfeller F R, Bapst U. Wireless in-house data communication via diffuse infrared radiation [J]. Proceedings of the IEEE, 1979, 67(11):1474–1486.
44. Barry J R, Kahn J M, Krause W J, et al. Simulation of multipath impulse response for indoor wireless optical channels [J]. IEEE Journal on Selected Areas in Communications, 1993, 11(3): 367–378.
45. S. Rodríguez, R. Pérez-Jiménez, F. J. López-Hernández, et al. Reflection model for calculation of the impulse response on IR-wireless indoor channels using ray-tracing algorithm [J]. Microwave and Optical Technology Letters, 2002, 32(4):296–300
46. S. Rodríguez, R. Pérez-Jiménez, O. González, et al. Concentrator and lens models for calculating the impulse response on IR-wireless indoor channels using a ray-tracing algorithm [J]. Microwave and Optical Technology Letters, 2003, 36(4):262–267.
47. Chi Nan, Lu Xingyu, Wang Can, et al. High-Speed Visible Light Communication Based on LED [J]. Chinese Journal of Lasers, 44(03): 7–18.

48. Chi N, Wang Y Q, Wang Y G, et al. Ultra-high-speed single red-green-blue light-emitting diode-based visible light communication system utilizing advanced modulation formats [J]. Chinese Optics Letters, 2014, 12(01):26–35.
49. Sun Huayan, Zhang Laixian, Zhao Yanzhong, et al. Progress of Free-Space Optical Communication Technology Based on Modulation Retro-Reflector [J]. Laser & Optoelectronics Progress, 2013, 50(04): 33–41.
50. Ke Xizheng, Lei Sichen. Modulation System of LED Visible Light Communication Uplink and Its Establishment Method [P]. China, CN201510056018.8, 2015-02-02.
51. Zhang Huan. Visible Light Full-Duplex Communication System Design [D]. [Master dissertation]. Xi'an University of Technology, 2018.
52. Yang Yufeng, Jiang Mingzheng, Zhang Ying, et al. Design of Full Duplex Visible Light Communication System Based on Single Light Source [J]. Laser & Optoelectronics Progress, 2019, 56(01): 78–86.
53. Zhang Ningning, Wu Shengju, Ling Xiaona. Study on Vibration Characteristics of Three Laminated Composite Piezoelectric Bending Transducer [J]. Piezoelectrics & Acoustooptics, 2015, 37(4): 620–625.
54. A R. Ndjiongue, H C. Ferreira, Jian Song, et al. Hybrid PLC-VLC channel model and spectral estimation using a nonparametric approach [J]. Transactions on Emerging Telecommunications Technologies, 2017, 28(12):168–180.
55. Lin Y J, Latchman H A, Lee M, et al. A power line communication network infrastructure for the smart home [J]. IEEE Wireless Communications, 2008, 9(6):104–111.
56. Xu Jidong, Wang Lin, Yang Liwei, et al. Research Report on the Combination of Visible Light Communication and Power Line Technology [R]. Beijing: China Communications Standards Association, 2014.
57. ITU-T Recommendation. Genenic home network transport architecture G. 9970 [S]. 2009.
58. Lin Yilei, Yu Jiangtao, Gu Youjia. Integrated communication system design of Li-Fi over powerline [J]. Energy Engineering, 2015, 6(06): 27–37.
59. Wang Yi, Xu Zhenzhao, Wang Nan. Combination of Power Line Communication and Visible Light Communication and Its Mobile Application [J]. Telecommunications Information, 2019, 6(11): 5–10.
60. Nlom Stephane M, Ndjiongue Alain R, Ouahada Khmaies. Cascaded PLC-VLC Channel: An Indoor Measurements Campaign [J]. IEEE Access, 2018, 25(6):230–239.
61. Ye Tian, Chang Jun Ahn. Performance Enhancement Using Unitary Frequency Modulation with PLC and VLC Cooperation [J]. International Journal of Communications, 2017, 39(11): 34–38.
62. GAO Shuli. A Dissertation Submitted to PLA Information Engineering University for the Degree of Master of Engineering [D]. [Master dissertation]. PLA University of Information Engineering, 2017.
63. Sun B, Ding W, Yan Y, et al. PLC and VLC integrated system for in-coach broadband video service [J]. Video Engineering, 2015, 39(11):63–66.
64. Yang Liwei, Hou Cong, Cheng Qiang. Research on Fusion Technology of Visible Light Communication and Power Line [J]. Telecommunications Network Technology, 2014, 2(02): 58–61.
65. Yang Xinhua. Modeling and Experimental Study of Visible Light Communication System [D]. [Ph. D. dissertation]. Jilin University, 2015.
66. Gao Shuli, Zhang Jian, Yang Hui. A reference channel model of indoor integrated power line communication and VLC systems [J]. Computer Engineering & Science, 2018, 40(06): 1062–1066.
67. Ma Q, Li C, Zhang Y, et al. Low cost lambda-tunable transmitter for indoor infrared wireless communication system [J]. Optics Communications: A Journal Devoted to the Rapid Publication of Short Contributions in the Field of Optics and Interaction of Light with Matter, 2019, 44(8):60–63.

68. Yan Chao, Xu Yunhao, Xue Leibing, et al. Heterogeneous indoor wireless access network based on collaborative VLC+WiFi [J]. Optical Communication Technology, 2016, 40(09): 41–43.
69. Shao S, Khreishah A, Ayyash M, et al. Design and Analysis of a Visible-Light-Communication Enhanced WiFi System [J]. IEEE/OSA Journal of Optical Communications and Networking, 2015, 7(10):960–973.
70. Z. Du, C. Wang, Y. Sun, et al. Context-aware indoor VLC/RF heterogeneous network selection: reinforcement learning with knowledge transfer [J]. IEEE Access, 2018, 6(07):33275–33284.
71. Pu Xin. Downlink Resource Allocation Scheme for Indoor VLC-WiFi Heterogeneous Networks [D]. [Master dissertation]. Chongqing University of Posts and Telecommunications, 2021.
72. Mohanad O., Anas M. S., Salam A. Z., et al. Joint optimization of power allocation and load balancing for hybrid VLC/RF networks [J]. IEEE/OSA Journal of Optical Communications and Networking, 2019, 10(5):553–562.
73. Wu Lan. Research on User Access and Resource Allocation for Indoor Heterogeneous Visible Light Communication Networks [D]. [Master dissertation]. Chongqing University of Posts and Telecommunications, 2019.
74. Zhu B. The Whole Monitor to the Performance of Optical Data Stream in All-optical Network [J]. Journal of Electronics & Information Technology, 2009, 31(1):206–209.
75. Carvalho J A R P D, Marques N, Veiga H, et al. Extended Field Performance Evaluation of a Gbps FSO Link [J]. e-Business and Telecommunications, 2012, 222(1):439–446.
76. Liu Yang. Research on the integration technology of LED visible light communication network and Ethernet [D]. [Master dissertation]. Heilongjiang University, 2019.
77. Zhao Zhizeng, Feng Chunpeng. Design of Ethernet interface to multi serial port transmission module based on STM32 [J]. Shanxi Electronic Technology, 2016, 12(2): 30–44.
78. Zhang Zhongyang, Tao Naibin. Ethernet interface design of embedded system based on ENC28J60 [J]. Journal of Henan University of Science and Technology: Natural Science, 2007, 28(3): 49–53.
79. Sun Zhaohui. Research and Implementation of Computer Remote Control Based on Trojan Horse [D]. [Master dissertation]. Tongji University, 2007.
80. Shen Jia. E3G Technology—The Core Technology and the Standardization of 3GPP LTE and 3GPP2 AIE [J]. World Telecommunication, 2006, 7(04): 29–46.
81. Yong-Hwan Son, Sung-Chan An, Hyun-Seung Kim, et al. Visible light wireless transmission based on optical access network using white light-emitting diode and electroabsorption transceiver [J]. Microwave and Optical Technology Letters, 2010, 52(04):790–793.
82. Wang Fang, Wang Zhaocheng, Qian Chen, et al. Efficient vertical handover scheme for heterogeneous VLC-RF systems [J]. IEEE/OSA Journal of Optical Communications and Networking, 2015, 7(12):1172–1180.
83. Liang Shufei, Zhang Yuexia, Fan Bo, et al. Multi-Attribute Vertical Handover Decision-Making Algorithm in a Hybrid VLC-Femto System [J]. IEEE Communications Letters, 2017, 21(7): 1521–1524.
84. W Huang, Y Yang, M Chen, et al. Wireless Network Optimization for Federated Learning with Model Compression in Hybrid VLC/RF Systems [J]. Entropy, 2021, 23(11):1413.
85. Zu Lijia. Research on Virtual Cell and Resource Management Technology of Visible Light Communication Network [D]. [Master dissertation]. Beijing University of Posts and Telecommunications, 2020.
86. Zhang Xinyue. Indoor Visible Light Communication Network Architecture and Resource Allocation [D]. [Master dissertation]. Beijing University of Posts and Telecommunications, 2014.

Optical-Camera Communication

36

Contents

Introduction 1384
Domestic and Foreign Research Progress 1385
System Model 1389
Working Principle of CMOS Image Sensors in Optical-Camera Communication 1390
Design of Data Frames for the OCC Sending End 1392
Design of the Demodulation Algorithm 1393
Selection of ROI 1393
Distortion Correction for Visible-Light Imaging 1393
Improved Center/Surround Retinex Image-Processing Algorithm 1399
Decoding Data 1404
Experimental Results and Analysis 1405
Bit-Error Rate Testing 1406
Analysis of Experimental Results 1408
Conclusion 1409
References 1410

Abstract

Optical-camera communication (OCC), a rapidly developing and emerging optical wireless imaging communication technology, has broad application prospects, owing to its low cost and highly popular optical-camera receivers. However, the received images have problems, such as perspective distortion, tilt distortion, halo effects, and optical noise, making threshold detection in image processing difficult. Implementing an OCC system that improves the data rate and transmission distance remains a problem. An image-optimization demodulation scheme is proposed. In the decoding process, the region of interest (ROI) is first used to extract the image matrix to reduce the impact of halo effects. Distortions in the image must be detected and corrected. We designed an improved center/surround retinex image-enhancement scheme, based on fast-scale adaptive bilateral texture filtering, which can efficiently eliminate the fluctuation of grayscale values both globally and locally. Subsequently, a simple threshold assessment is

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_36

performed. The stripe width of the received pixel is used for symbol judgment to complete subsequent demodulation and decoding; then, the original transmission data are obtained. In an experimental analysis, the improved center/surround retinex algorithm proposed in this study solved the problem of fluctuating grayscale values. Compared with other algorithms, it had the highest decoding efficiency, whether with an increase in communication rate from 3.5 kbps to 6.5 kbps or a change in communication distance. It had an average decrease in bit-error rate of about one order of magnitude and the best decoding performance.

Keywords

Optical camera communication · Halo effect · Distortion correction · Center/surround retinex

Introduction

Visible-light communication (VLC) technology, one of the most promising wireless communication technologies in the 6G blueprint, has attracted widespread attention, owing to its rich spectral resources, fast transmission speed, ultra-high bandwidth, excellent security, and lack of electromagnetic radiation [1]. Currently, VLC, as an integrated green lighting and communication function, plays an important role in short-range communication [2]. Optical-camera communication (OCC), an important branch of low-power and broad-spectrum VLC, has become more popular, owing to the widespread development of complementary metal-oxide semiconductor (CMOS) cameras at the receiving end.

Optical-camera communication encodes and modulates the transmission data to be input at the sending end; the coding prevents as much as possible the occurrence of a stroboscopic light-emitting diode (LED) light-source effect [3]. If people's eyes are exposed to a stroboscopic environment for a long time, it can affect their mood, causing boredom, anxiety, etc. It may result in eye fatigue, distraction, and in serious cases, migraine headaches, blurred vision, epilepsy, and possibly neurological diseases.

The CMOS optical camera at the receiving end is equipped with a high-resolution CMOS image sensor [4], and its rolling-shutter effect greatly improves the data rate, compared to the global-shutter effect. The optical signal at the sending end is recorded as bright and dark stripes representing "1" and "0" [5, 6].

Although the shutter mode improves the data-transmission speed, it introduces a series of problems. One problem is that the data between two imaging frames cannot be captured. In the design process, it should be ensured that each captured image frame contains at least one complete data packet. Another issue is the correction of image perspective and tilt distortion caused by the offset of the receiving angle at the receiving end and jitter during shooting. A third issue is the halo effect, which is caused by the overflow of charges from saturated pixels to adjacent pixels. The light-

overflow effect and optical noise make threshold decisions in image processing difficult, which can reduce the decoding efficiency.

To solve the above problems, we designed correction methods for the tilt and perspective distortion caused by unstable receiver equipment and different angles. Moreover, we proposed an improved center/surround retinex algorithm based on fast-scale adaptive bilateral smooth filtering to eliminate the global and local fluctuations of gray values. A simple threshold decision can efficiently decode the original data. We tested the bit-error rate under different transmission rates and distances and concluded that our algorithm could efficiently decode the sent data in both cases.

Domestic and Foreign Research Progress

In 2000, Japanese scholar Tanaka et al. [7] proposed the use of white LEDs to simultaneously provide indoor lighting and wireless data transmission using light, designed a wireless home-link communication system, and simulated and analyzed the system. In 2006, Ding and Ke [8] from Xi'an University of Technology first introduced the concept of visible-light communication into China. With the continuous advances in key technologies in the field of visible-light communication, breakthroughs have been made in many applications.

In 2011, the IEEE released the IEEE Standard 802.15.7-2011 visible-light communication standard to ensure standardized research for VLC systems. Its full name was the visible-light short-range optical wireless communication standard, which was the first standard related to VLC. Although the IEEE 802.15.7 standard originally developed by the IEEE did not include optical wireless communication based on optical cameras, it was included in the 2014 IEEE 802.15.7r1 revised standard. After OCC was included in the revised standard, the standard was managed by a TG7m organization composed of academic researchers from multiple countries, such as China, the United States, and South Korea. The initial draft of IEEE 802.15.7-2018 was completed in 2018 [9].

In 2012, Danakis et al. [10] proposed a CMOS-camera data-reception scheme using visible-light communication, using a cubic-polynomial fitting method to distinguish the LED status, achieving a data rate far higher than the camera frame rate for visible-light imaging communication systems.

In 2013, Takai et al. [11] proposed a fast LED detection technology and integrated communication pixels, ordinary image pixels, and fast LED detection-technology circuits in optical-communication image sensors to produce a special CMOS image sensor. They designed a new optical wireless communication system using LEDs and CMOS technology for automotive applications.

In 2014, Chen et al. [12] proposed and demonstrated a novel filterless VLC system based on RGB light-emitting diodes and mobile-phone cameras. The system utilized a feature-matching method based on a scale-invariant feature-transformation algorithm for the received grayscale images. The experiment showed that the proposed method was simple and reduced the computational complexity.

In 2015, Chow et al. [13] proposed a link for visible-light imaging communication using CMOS cameras, which had a data rate higher than the frame rate of CMOS image sensors. At the receiving end, histogram equalization and Sobel filters were used to enhance the VLC signal performance, which alleviated the impact of halo effects. However, a significant bit-error rate occurred when the communication rate increased.

In the same year, Chow et al. [14] proposed using second-order polynomial fitting at the receiving end of visible-light imaging communication systems to reduce the halo effect of CMOS sensors. They used histogram equalization and Sobel filters to enhance the VLC signal performance, and third-order polynomial fitting to determine the threshold. The experimental results showed that, under different illuminances, the error rate could be improved by approximately an order of magnitude.

In 2016, Liu et al. [15] proposed and experimented with three threshold schemes, namely third-order polynomial curve fitting, iterative schemes, and fast-adaptive schemes, to evaluate them. In the same year, Liu et al. [16] proposed a visible-light imaging communication scheme using devices with LED and CMOS sensors for optical encrypted transmissions. They proposed and demonstrated the use of the Otsu threshold to define the data logic in shutter mode, and reduced the problem of fluctuating grayscale values using smoothing schemes.

In the same year, Liu et al. [17] proposed and demonstrated the use of a second-order polynomial-enhanced extinction-ratio scheme and an iterative and improved fast adaptive-threshold scheme to demodulate the obtained shutter mode. The experimental results showed that by using this enhancement scheme, the error rate could be improved by two orders of magnitude.

In 2016, Liang et al. [18] studied the performance of visible-light imaging communication systems when light sources moved to different parts of an image frame, and proposed and demonstrated the use of region-growth algorithms to track light sources. The extinction-ratio problem was reduced through selecting a grayscale-column pixel matrix, second-order polynomial smoothing, and Sobel filtering.

In the same year, Liang et al. [19] studied the effective thresholding problem in the experimental demodulation process of visible-light imaging communication systems. They proposed and demonstrated the maximum-entropy and minimum cross-entropy threshold algorithms. The demodulation error rates of the two algorithms were equivalent; however, the maximum-entropy threshold algorithm had a lower latency and was more advantageous.

In 2016, Li et al. [20] proposed a digital image-processing scheme for visible-light imaging communication using mobile-phone cameras as receivers. The scheme utilized grayscale conversion to remove low-grayscale and LED light spots and histogram equalization and high-pass filters to process the image. The resulting bright and dark stripes were clearer and easier to decode. The experimental results demonstrated that the VLC system could resist interference and extend its transmission distance.

In 2017, Dai et al. [21] proposed a non-line-of-sight channel communication scheme to solve the halo effect caused by light sources in optical-camera

communication and the communication-discontinuity problem between multiple frames. The experimental results showed that the synchronization accuracy and communication speed were improved through high-precision synchronization algorithms and multi-column signal accumulation.

In the same year, Chen et al. [22] proposed an effective demodulation scheme to synchronize and demodulate the shutter mode of visible-light imaging communication systems. First, linear interpolation was used to increase the number of effective sampling points, and then a new threshold scheme using extreme-value averaging was proposed. The experimental results showed that the proposed scheme significantly improved the processing latency and reduced the bit-error rate.

In the same year, Wang et al. [23] proposed a long-distance non-direct-view link communication system that used improved rolling-shutter mode demodulation algorithms, such as background-compensated floodlight mitigation, extinction-ratio enhancement, and the Bradley adaptive threshold, which could receive and demodulate 1.5-m non-direct-view link communication under a low illumination of 145 lx.

In the same year, Zhang et al. [24] proposed a scheme to enhance the performance of visible-light imaging communication using a moving-exponential averaging algorithm. This scheme first sorted the elements of each row of the grayscale values of the pixel matrix in descending order and then selected an appropriate pixel array to reduce random noise interference. The moving-exponential averaging algorithm reduced the impact of extinction-ratio fluctuations on the decoding performance.

In 2018, Lee et al. [25] proposed a method for suppressing the noise interference in visible-light imaging communication systems. The proposed scheme used "signal averaging" to reduce additive white Gaussian noise, and reduced background noise through an adaptive window threshold and normalization. Experiments showed that this scheme improved the transmission performance of visible-light imaging communication systems.

In the same year, Shi et al. [26] proposed using a dynamic-column matrix-selection algorithm to avoid halo effects and experimented with optical-camera communication that supported user mobility. The results showed that the bit-error rate performance was inversely proportional to the movement speed and lateral distance, and decreased with a corresponding increase. In the same year, Guan et al. [27] proposed the use of contrast-constrained adaptive histogram equalization to extend the transmission distance and reduce the effect of background noise.

In the same year, Sun [28] investigated the visible-light communication system of smartphones, designed a signal-reception scheme based on mobile cameras and mobile light sensors, and determined the practical value of the mobile-camera scheme. This system effectively reduced the hardware requirements of traditional VLC systems for docking with the receiving users. In the same year, Zhang et al. [29] proposed a threshold algorithm based on stripe-boundary pixels to improve the performance of visible-light imaging communication systems. The results showed that this algorithm had advantages in terms of transmission distance and data rate, compared with traditional third-order-polynomial fitting schemes.

In 2019, Zhang et al. [30] proposed three threshold algorithms to improve the performance of VLC systems based on the CMOS-camera shutter effect. The

threshold schemes were based on equal-weight, fractional-weight, and ratio-weight moving averages. Compared with traditional threshold schemes based on polynomial fitting, the proposed threshold scheme had significant advantages in terms of both system performance and computational complexity. The threshold scheme based on the fractional-weight moving average was the best.

In the same year, Li et al. [31] proposed an efficient thresholding method to increase the transmission distance in optical-camera communication. This method utilized a nonlinear transformation of the average grayscale values of each column to increase the contrast between the light and dark bands. The experimental results demonstrated that this method was superior to polynomial regression.

In the same year, Zhang et al. [32] proposed an information-decoding scheme for visible-light imaging communication systems that effectively suppressed the halo effect of received images by utilizing synchronous dynamic-compensation technology and introducing histogram equalization, least-squares fitting, and other technologies. The results demonstrated that this decoding scheme effectively improved the reliability of information transmission.

In the same year, Zhong et al. [33] designed an algorithm based on back-propagation neural networks to decode images received by lens-less optical-camera communications. This algorithm effectively recognized blurred image patterns and used an adaptive training-sequence adjustment mechanism to adapt to time-varying environments. Simulation results showed that the proposed image-decoding algorithm had a good bit-error rate.

In the same year, Zhu et al. [34] proposed a frame-synchronization mechanism for a rolling shutter based on the imaging principle of optical cameras, applied visible-light imaging communication to directional transmission characteristics, and designed a sensor-node debugging information-security transmission scheme based on optical-camera communication.

In 2020, Liu et al. [35] proposed an adaptive contrast-enhancement algorithm to solve the problem of high-extinction-ratio images in visible-light imaging communication systems. This scheme was suitable for high-fluctuation data modes; however, when the transmission distance increased, the decoding error rate exceeded the forward error-correction threshold limit of 7%. In the same year, Meng et al. [36] used a feature grayscale matrix to alleviate the impact of the halo effect but did not fundamentally eliminate the grayscale-value fluctuation problem.

In 2021, Li et al. [37] proposed a novel demodulation scheme for visible-light imaging communication systems that utilized an adaptive image-enhancement algorithm based on dilation and corrosion to solve the problem of fluctuations in grayscale values. However, it occasionally deformed pixel objects and could lead to information loss. In the same year, Bai et al. [38] designed a new data-frame structure suitable for optical-camera communication links, effectively solving the data-loss problem in optical-camera communication links caused by asynchronous receiving and sending ends and camera-frame spacing.

In 2022, Wu et al. [39] proposed a multilevel modulation scheme using a pilot-assisted K-means algorithm that obtained the stripe-propagation state in optical

channels under different environments and improved the data rate and performance of visible-light imaging communication.

In the same year, Dong et al. [40] proposed a simple channel-equalization method that utilized the spatiotemporal relationship of spatial-delay pulse-width modulation waveforms transmitted by multiple light sources to effectively alleviate the uncertain sampling problem caused by the asynchronous nature of optical-camera communication. They analyzed the equalization error rates caused by undersampling and overexposure under different duty cycles and exposure times.

In the same year, Zhang et al. [41] proposed introducing visible-light imaging communication into subway-train positioning, using LED light sources in tunnels as the information-sending end of the positioning system and optical cameras as the information-receiving end. Visible-light imaging communication technology was used to obtain the coordinates and shape information of the LED light sources to determine precise train positions.

In the same year, Chen et al. [42] proposed an exponentially-weighted filtering algorithm to suppress the interference of grayscale fluctuations, ensuring the influence of adjacent pixel rows. This algorithm, combined with reference-stripe judgment, rapidly demodulated and decoded the data of visible-light imaging communication systems, reduced the problem of high error rates caused by grayscale fluctuations in optical-camera communication, and had lower algorithm complexity.

In the same year, Jiang et al. [43] proposed a pilot-based strip-area-estimation decoding algorithm and designed pixel-rearrangement schemes and stripe-separation decision-making methods that improved the decoding performance under receiver sway or position shifts in visible-light imaging communication. The effectiveness of the proposed decoding algorithm was experimentally verified.

In 2023, Zhang et al. [44] proposed a train-positioning algorithm based on the combination of a backpropagation (BP) neural network and visible-light imaging communication. This method used an LED light source in the tunnel as the communication transmitter, introduced feature variables in the light-stripe-code modulation process, and used a BP neural network to classify and recognize the information sent by the light source. This increased the communication distance and improved the timing and accuracy of train positioning.

System Model

The system model of optical-camera communication is shown in Fig. 36.1. The computer inputs the data to be transmitted. The signal light source at the sending end adopts Manchester encoding, which can continuously provide up to two identical bits, preventing the stroboscopy problem caused by a large number of consecutive 1 or 0 bits. Using binary amplitude on–off keying (OOK) modulation, the 01 code of the transmitted data is represented by a unipolar non-return-to-zero code. After the data are encoded and modulated, the microcontroller controls the LED driver circuit to send the data to the LED light source for transmission.

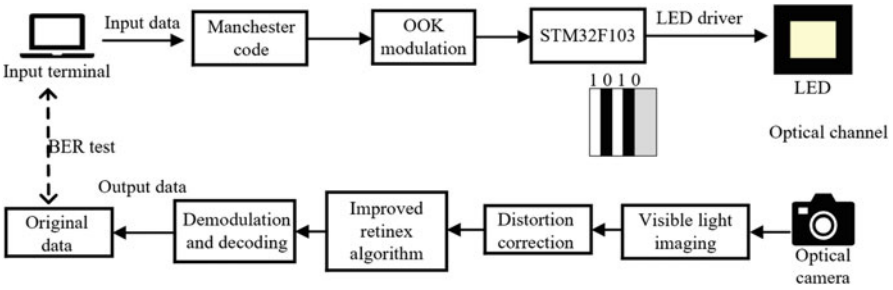


Fig. 36.1 Overall block diagram of an optical-camera communication system

After transmission through the optical channel, it reaches the receiving end and is received by an optical camera with a CMOS image sensor, forming an image with light and dark stripes. The received image is optimized and processed, including correcting for distortion and eliminating overall and local grayscale fluctuations in the image, to achieve efficient demodulation and decoding. Finally, the original transmission data are obtained.

Working Principle of CMOS Image Sensors in Optical-Camera Communication

Optical-camera communication uses optical cameras as receivers of optical signals, and is a new type of visible-light wireless communication. The light source at the transmission end performs the dual tasks of lighting and communication.

CMOS optical cameras have become one of the most promising receiving platforms for optical-camera communication because CMOS image sensors use a rolling-shutter exposure mode. Specifically, the light signals sent by LED light sources are captured by the camera in the form of images; the signal-capturing process is shown in Fig. 36.2. Compared with the global-exposure mode of CCD imaging, the rolling-shutter exposure mode improves the transmission rate of optical-camera communication [45].

The CMOS image sensor performs row-by-row exposure under shutter-mode control. When the LED is turned on, the pixels are exposed to form bright stripes; when the LED is turned off, the pixels form dark stripes. The received light signal is converted into an electrical signal and recorded to obtain a strip image with alternating light and dark stripes. The transmitted data can be obtained by analyzing the light and dark stripes [45].

The frame rate R_i of a camera is generally expressed as the reciprocal of the sum of the frame-reading time T_R and the interframe interval T_G, as shown in Eq. (36.1).

$$R_i = \frac{1}{T_R + T_G} \tag{36.1}$$

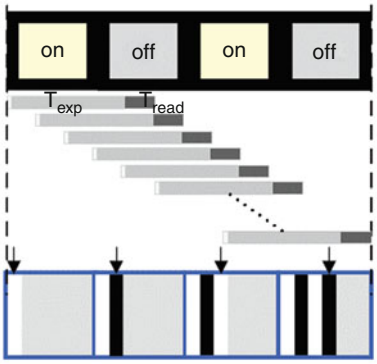


Fig. 36.2 Rolling-shutter exposure mode of a CMOS image sensor

The frame-readout time is $T_R = T_{\text{exp}} + (Y - 1) \times T_{\text{read}}$, where T_{exp} is the pixel-row exposure time of the CMOS image sensor, T_{read} is the reading time of the CMOS image sensor, and Y is the total number of pixel rows in the image.

To ensure correct data reception, it is necessary to consider the problem of matching the visible-light transmission rate and the shooting frame rate of the camera, as shown in Eq. (36.2). This means that the transmitted data rate R_r should simultaneously be higher than the shooting frame rate R_i of the camera and lower than the row pixel-scanning rate R_s of the CMOS image sensor. The shutter time of the CMOS image sensor determines the rate range of the visible-light information.

$$R_i \leq R_r \leq R_s \tag{36.2}$$

The row pixel-scanning rate R_s of CMOS image sensors can be estimated from the total number of pixel rows in the image and the frame rate of the camera; it is represented as

$$R_s = Y \times R_i \approx Y \times \frac{1}{T_R} \tag{36.3}$$

The frequency of LED flashing must be set within an appropriate range, ensuring that the frequency exceeds the human-eye limit and that normal lighting functions are achieved, as well as correct reception by CMOS image sensors. The maximum allowed data-transmission rate can be expressed as

$$R_{r\,\max} = \frac{1}{t_{\min}} \tag{36.4}$$

where t represents the exposure time of the camera in seconds and $t_{\min}$ represents the shortest exposure time. Although the parameter values of optical cameras vary among different models, the shorter the exposure time, the higher the system communication rate.

After the data rate R_r transmitted by the system is determined, the bit resolution of an image frame can be further determined at a specific receiving end. The bit resolution is expressed as

$$r = \frac{R_i \times Y}{R_r} \tag{36.5}$$

The number of information bits that can be carried in an image is represented as

$$N_r = \frac{Y}{r} \approx \frac{R_r}{R_i} \tag{36.6}$$

It can be seen that the number of bits transmitted in an image frame is closely related to the transmission rate of the system and the frame rate of the camera.

Design of Data Frames for the OCC Sending End

The data-frame design should not only ensure the correct reception of data information but also consider the issue of synchronous reception at the receiving end. In this experiment, the frame header of the data frame adopts an identifier with a length of eight bits, designed as 01111101, to ensure that the receiving end recognizes the frame header to achieve a synchronous reception. This frame flag can serve as both the start flag of one data frame and the end flag of the previous data frame, distinguishing between the two data frames. The 01 code in the frame identifier can also serve as a reference for the subsequent decoding.

Then, an eight-bit sequence identifier is designed to ensure that the receiving end decodes the data and reorders them correctly. A data frame can transmit 16 or more bits of data. In the data-frame design, the sequential identifier and data were encoded using Manchester encoding, as shown in Fig. 36.3.

Equation (36.1) shows a frame gap T_G between two consecutive images captured by the camera, $T_G = \frac{1}{R_i} - T_R$, which can lead to incomplete data information at the receiving end. The sending rate of the LED can be adjusted to ensure that at least two identical data frames are recorded in one frame of the image at the receiving end.

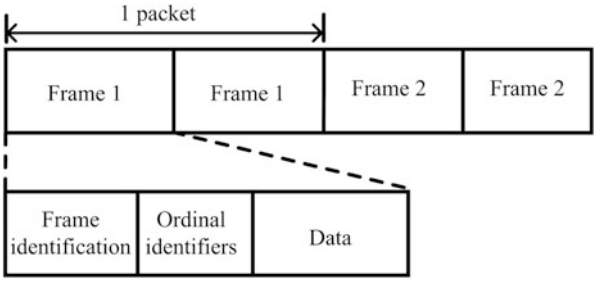


Fig. 36.3 One packet containing two data frames

Including two data frames in one packet can solve the problem of data loss caused by inter-frame intervals, thereby ensuring that complete data are obtained when decoding the received images.

Design of the Demodulation Algorithm

There are several basic problems in experimental research on optical-camera communication. Problems with perspective and tilt distortions in imaging are caused by the shooting angle and shaking during the shooting. Because of the interference of ambient light, the received signal is polluted by additive white Gaussian noise. Severe fluctuations affect grayscale image values. These issues cause serious errors in the subsequent decoding, and the receiving end will not receive the correct transmission data.

To solve the problems with receiving in optical-camera communication, this study designed corresponding solutions. Corresponding fast-correction methods were designed for the two types of distortion problem. An improved center/surround retinex algorithm was designed to address the interference of light noise in the environment and the fluctuation of grayscale values in the images. Changing the kernel function of the center/surround retinex algorithm enabled it to solve the problems of overall and local grayscale fluctuations in the received image, thereby ensuring correct decoding of the transmitted data stream. Because optical-camera communication uses simplex communication, reducing the error rate is a key factor in achieving correct communication.

Selection of ROI

To reduce the halo effect and improve the timeliness of subsequent decoding, the region of interest (ROI) was selected according to the method described in [46], as shown in Fig. 36.4. By selecting an image matrix with a smaller halo effect, the grayscale feature curve shown in Fig. 36.5 can be obtained. However, there is still a serious fluctuation in the grayscale values in the selected ROI, which affects the subsequent decoding.

Distortion Correction for Visible-Light Imaging

During the experiment, the receiving end may be in different situations. In one situation, different shooting angles at the receiving end can change the stripe width of the image, which is a perspective distortion. In another situation, the shaking of the device at the receiving end can cause the image to have skewed stripes, which is known as tilt distortion. In both cases, the image at the receiving end is distorted and cannot be recognized. This seriously affects the visible-light imaging communication and is unacceptable.

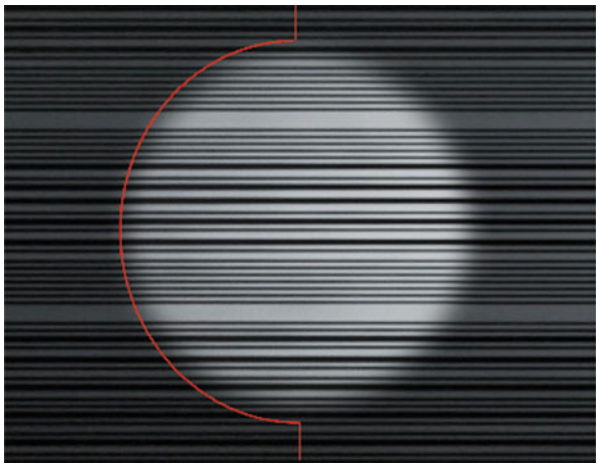


Fig. 36.4 Selection of ROI image matrix with a low halo effect

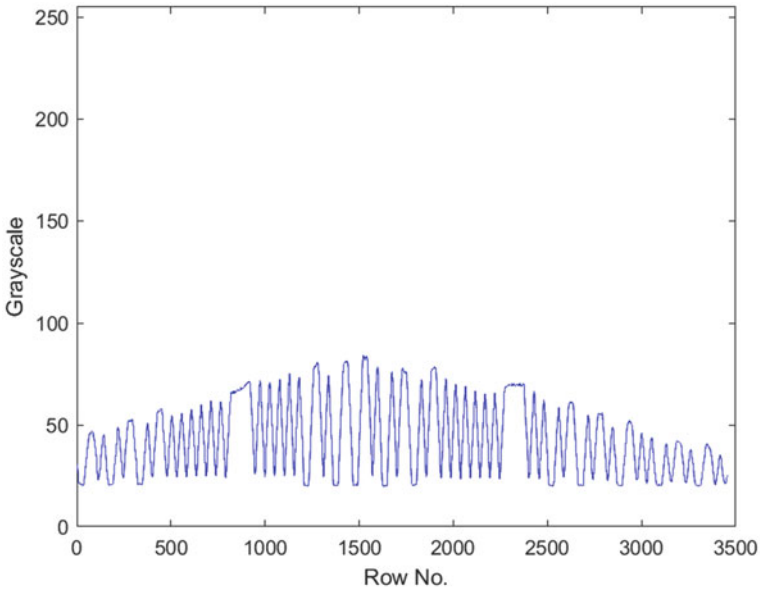


Fig. 36.5 Grayscale-feature curve of ROI region

Therefore, correcting the distortion of the received image is crucial for ensuring correct reception at the receiving end of an OCC system. Considering the different characteristics of these two types of distortion, two different correction methods have been proposed.

Correction of Perspective Distortions

During the experiment, there may be situations in which the receiving end does not directly face the light source. For example, when shooting from the left or right, the received image features may differ, which can lead to perspective distortion.

The imaging results obtained when shooting from the right side of the light source are shown in Fig. 36.6. The dashed line represents the original image size, and the solid line represents the distorted shape. The perspective distortion is due to the imaging characteristics of the optical-camera equipment itself, which is "near large, far small."

The specific phenomenon of perspective distortion is that the stripe width of the receiving-end image on the side near the camera is significantly greater than that on the side away from the camera. Perspective distortion changes the original width of the image stripes. Eliminating perspective distortion restores an image to its actual size. A perspective-transformation method was proposed to solve the problem of perspective distortion caused by the camera.

Perspective transformation is the process of projecting an image onto a new visual plane for correction [47]. The process involves finding the four corners of the image and converting them into equal widths, based on the perspective transformation. In this study, four preset corner points were used to resolve the perspective distortion caused by the imaging characteristics of the camera. The advantage of this perspective-transformation correction method is that it has lower computational complexity and faster processing, making it a very effective transformation method.

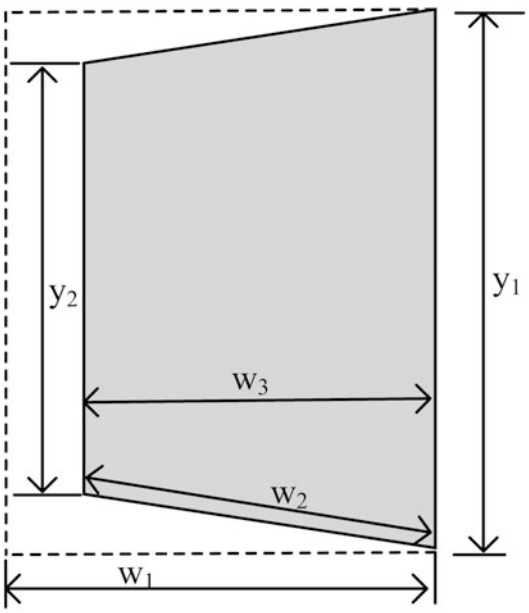


Fig. 36.6 Schematic diagram of perspective distortion

Based on the characteristics of visible-light imaging perspective distortion, this study vertically divided the distorted pixels into three small pixel parts. These three parts included the distortion on the side near the camera, the distortion far from the camera, and the portion in the middle. The segmented small pixels were stretched separately and recombined to obtain the original image.

Segmentation correction can obtain the corner coordinates of the segmented image, thereby accelerating the subsequent calculation process. The specific correction process is as follows: The perspective transformation projects the image onto a new viewing plane. The transformation formula is shown in Eq. (36.7), where (x, y, z) are the coordinates of the original image pixel, (X, Y, Z) are the coordinates of the transformed image pixel, and A is the perspective-transformation matrix.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{36.7}$$

The image is then converted into a two-dimensional image. $Z = 1$ and $a_{33} = 1$. The coordinates of the transformed image can be obtained as

$$\begin{cases} a = \dfrac{X}{Z} = \dfrac{xa_{11} + ya_{21} + a_{31}}{xa_{13} + ya_{23} + a_{33}} \\ b = \dfrac{Y}{Z} = \dfrac{xa_{12} + ya_{22} + a_{32}}{xa_{13} + ya_{23} + a_{33}} \end{cases} \tag{36.8}$$

To correct a perspective distortion, four original coordinates and four destination coordinates must be known. The distorted image is divided into three uniform parts. After the segmentation, the pixel coordinates of the four corners of each part can be obtained. The pixel coordinates of the four corner points are $I(x_1, y_1)$, $I(x_2, y_2)$, $I(x_3, y_3)$, and $I(x_4, y_4)$. The destination coordinates are the known coordinates of the specified rectangle for the new view. The pixel coordinates of the four corner points are (a_1, b_1), (a_1, b_2), (a_2, b_1), and (a_2, b_2).

Subsequently, by substituting these eight pairs of coordinates into Eq. (36.8), the remaining eight parameter values in the parameter-transformation matrix A, excluding $a_{33} = 1$, can be calculated. Thus, the distorted image can be corrected to obtain a normal image, and further demodulation and decoding can be completed.

The perspective-distortion correction process is shown in Fig. 36.7. Figure 36.7a shows the image captured when turning the camera angle to the right, Fig. 36.7b shows the pixel coordinates of the four corner points obtained by segmenting and labeling the distorted-perspective image, and Fig. 36.7c shows the corrected image.

Correction of Tilt Distortions

At the receiving end, a tilt distortion may occur, which differs from a perspective distortion. When the receiving device is unstable and sways during the shooting

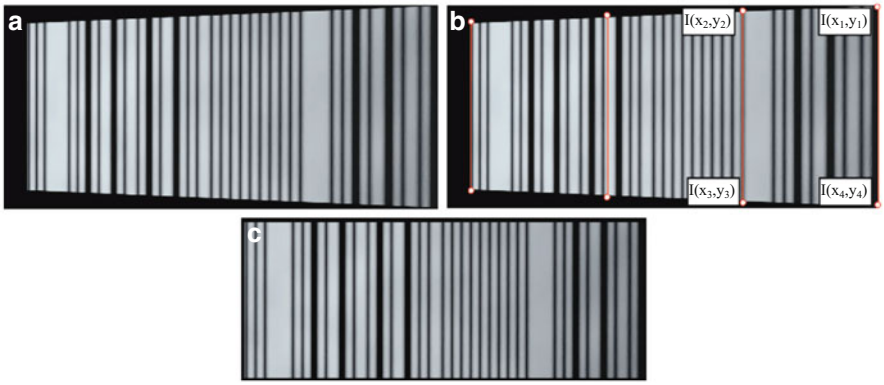


Fig. 36.7 Perspective-distortion correction process: (**a**) Picture taken from the right side, (**b**) marked segmented image, and (**c**) corrected image

process, it can cause the received image to tilt and shift, which is known as tilt distortion. As shown in Fig. 36.9a, the overall stripes in the image are tilted; however, the width of the stripes remains unchanged, which affects the ability to proceed with subsequent work.

According to the characteristics of visible-light imaging, the resulting image is a bright–dark-striped image. A random Hough transform (RHT) can be used to quickly correct the tilt problems caused by the shaking of the receiving optical camera. The Hough transform utilizes the transformation between the space in which the image is located and the parameter space for calculations [48]. The random Hough transform uses two points in Cartesian coordinates to obtain any parameter point in the parameter space; it has low computational complexity and high accuracy.

The random Hough transform is exceptionally well-suited for correcting tilt distortions, owing to its robust resistance to interference, small storage requirements, rapid straight-line detection capabilities, and precise correction of image-tilting issues.

The transformation relationship of the RHT is shown in Fig. 36.8, and the slope-intercept parameter equation for the straight line in the image is

$$f(x, y) = y - kx - b = 0 \tag{36.9}$$

where k represents the slope of the line and b represents the y-intercept. Taking k as the independent variable and b as a function, we obtain the expression in the parameter space:

$$b = -kx + y \tag{36.10}$$

The visible-light imaging tilt-correction process projects the tilted image onto a new visual plane. In the process of correcting tilted images, the angle size of the

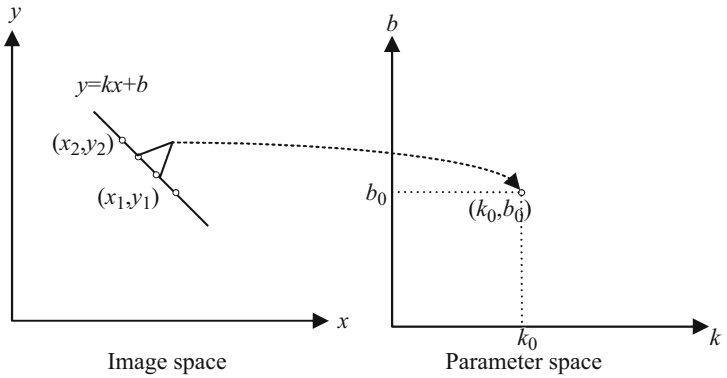


Fig. 36.8 RHT conversion relationship

image tilting can be determined, based on the frame-flag header, which improves the speed of correcting the tilted image. The tilt-correction process is as follows:

1. Because at least one data frame is present in the image, that is, at least one frame flag header, the position of the image-frame flag header is stored in dataset D.
2. Two pixel points, (x_1, y_1) and (x_2, y_2), are randomly selected from dataset D of the frame flag header.
3. From Eq. (36.10), the two pixel points, (x_1, y_1) and (x_2, y_2), selected in Step (2) can be mapped to a point (k, b) in the parameter space, and then the point (k, b) is placed in parameter set P.
4. After repeating steps (2) and (3) a few times, a parameter $y = k_1x + b_1$ in the image space can be obtained. At this point, a certain number of parameters have accumulated in parameter set P. By searching for these parameter points in P, the lines in the frame-flag header can be detected.
5. The image is corrected, based on the inclination angle of the straight line.

Figure 36.9 shows the process of correcting the tilted image angle. Figure 36.9a shows the tilted image, and Fig. 36.9b shows the image with the tilt corrected. However, the corrected image is incomplete, with missing pixels, as shown in Fig. 36.9b, and subsequent processing cannot be performed at this time.

In this study, a method was designed to fill in the missing pixels by scanning each column of pixels in the corrected image, based on the image characteristics. During this correction process, the size of the image pixels, namely, the width and height, were also calibrated; thus, the pixel size of the filled image remained unchanged, resulting in a complete image, as shown in Fig. 36.9c, which could be used for subsequent processing experiments.

When perspective or tilt distortions exist, the image at the receiving end cannot be recognized and subsequent work cannot be carried out. In this study, different correction methods were adopted, based on the different phenomena caused by these two types of distortion.

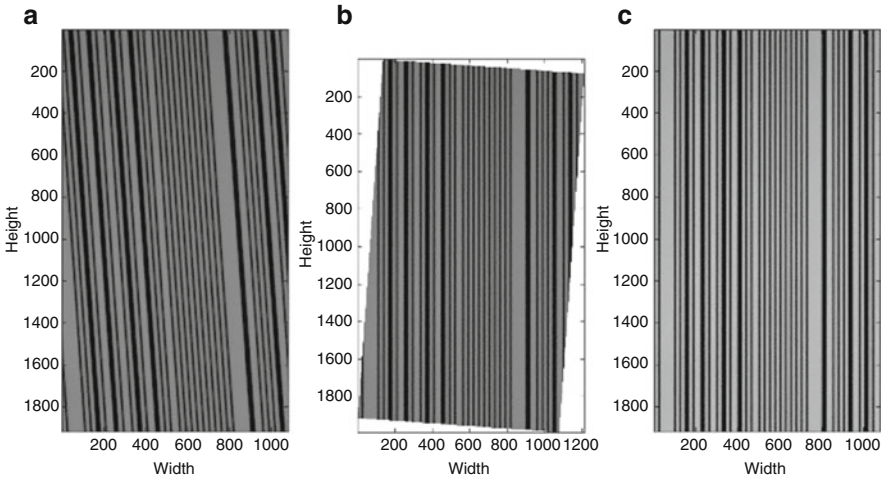


Fig. 36.9 Tilt-distortion correction process: (**a**) Tilted image, (**b**) corrected tilted image, and (**c**) complete image

Although a perspective transformation can be used to detect straight lines, in this design, it mainly addresses the perspective distortion caused by the shooting angle. The perspective-transformation method is used for correcting perspective distortion. If there is only tilt distortion caused by simple jitter, it is more suitable to use a random Hough transform because it has a lower computational complexity and higher speed.

Improved Center/Surround Retinex Image-Processing Algorithm

Center/Surround Retinex Theory

The retinex theory suggests that an image $S(x, y)$ is composed of an illumination component $L(x, y)$ and a reflection component $R(x, y)$ [49]. Its relationship is expressed in logarithmic form, as shown in Eq. (36.11). The illumination component represents the low-frequency part of the image and determines its dynamic range. The reflection component corresponds to the high-frequency part of the image and represents its essential characteristics.

The basic idea of the retinex theory in image enhancement is to remove the influence of the environment illumination component on the image and obtain the reflection component, which is the enhanced image.

$$\log S(x,y) = \log L(x,y) + \log R(x,y) \tag{36.11}$$

The real-number field formula after image enhancement is

$$\begin{aligned} r(x,y) &= \exp[\log R(x,y)] = \exp[\log S(x,y) - \log L(x,y)] \\ &= \exp\{\log S(x,y) - \log[G(x,y)^{*} S(x,y)]\} \end{aligned} \tag{36.12}$$

where $G(x, y)$ represents the surround function. The grayscale values of the pixels in an image are calculated by weighting the pixel values in the area centered on the selected pixel. The weight ratio is determined by the surround function $G(x, y)$. It is used as a filter function to calculate the illumination component in the image; a Gaussian function is generally used as the surround function. $*$ represents the convolution operation between the surround function and the original image. To expedite the computation during image processing, the convolution operation in the time domain was converted to a multiplication operation in the frequency domain.

Improved Center/Surround Retinex Algorithm

When the Gaussian function is used as the surround function to calculate the illumination component, the edge information in the image cannot be well preserved, resulting in unclear details at the image edge, which significantly affects the final image enhancement.

Accordingly, this study proposes an improved center/surround retinex algorithm to enhance the image of an optical-camera communication receiver. The specific implementation replaces the convolutional surround function with a fast scale-adaptive bilateral texture-smoothing filter. The algorithm implementation process is illustrated in Fig. 36.10. The algorithm is as follows:

1. A Gaussian function is used to extract the gradient information of the original image. The extracted gradient information is used as the information basis of the spatial-kernel box filter in the surround function to adaptively adjust the spatial-kernel width at each pixel.
2. A fast scale-adaptive bilateral texture-smoothing filter is convolved with the original image to obtain the illumination component.
3. The reflection component in the logarithmic domain is obtained from the original image and illumination components.

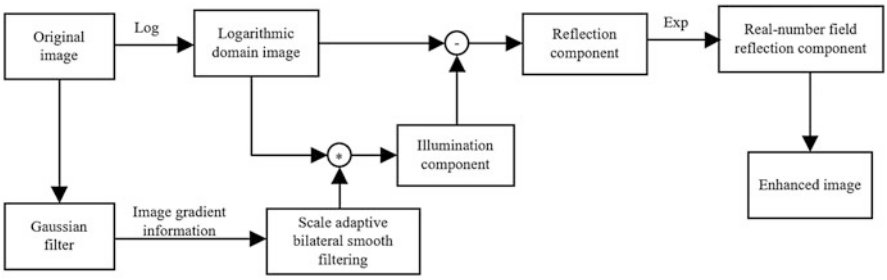


Fig. 36.10 Implementation process of the improved center/surround retinex algorithm

4. The reflection component is then converted into a real-number field and adjusted to obtain an enhanced image.

The processed image removes the influence of light noise in the environment, preserves the edge information of the image, and solves the problem of grayscale-value fluctuation to the maximum extent, both globally and locally.

An image can be decomposed into structural and textural components, in which the structural components are composed of smooth regions and significant geometric edges, whereas the texture is composed of small-scale changes. The non-uniformity of the texture components can lead to the loss of structural information, excessive smoothing, and blurring of edge information [50].

The bilateral filter uses a range kernel and spatial kernel to preserve the image-edge information while achieving a good smooth denoising effect; however, it cannot be directly used to smooth the rough texture while preserving the edge information. The scale-adaptive bilateral texture-smoothing filter can preserve the fine details and clarity of the structural components while smoothing out noise; it also clarifies the details at the edges.

The convolutional surround function $G(x, y)$ is a scale-adaptive bilateral texture-smoothing filter, and its gradient-based rules can adaptively adjust the spatial kernel at each pixel to obtain better filtering results. The range kernel provides the intensity differences in the image, preventing edge blurring and pixel mixing from both sides of the edge, which are beneficial for preserving the structure.

The formula for the scale-adaptive bilateral filter of the input image f is as follows:

$$S(i) = \frac{\sum_{j \in \Omega_i} \omega(j-i)\varphi(f(j)-f(i))f(j)}{\sum_{j \in \Omega_i} \omega(j-i)\varphi(f(j)-f(i))} \tag{36.13}$$

where $S(i)$ is the output value at pixel i, and $\omega(l)$ and $\varphi(t)$ are the spatial and range kernels, respectively.

The Gaussian range kernel is

$$\varphi(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{36.14}$$

where σ is the standard deviation of the Gaussian kernel, which determines the width of the range kernel. In the experiment, we found that when the data-transmission rate was higher, the value of σ should be correspondingly reduced to ensure that the structural edges are not blurry and the imaging is clear.

The innovation, compared with the classic bilateral filter, is that the spatial kernel uses a box filter. The window Ω_i of the spatial kernel is centered on pixel i; $|\Omega_i| = (2w_i + 1) * (2w_i + 1)$, where i is the pixel belonging to Ω_i, and w_i represents the proportion of the spatial kernel at pixel i. By adjusting the width of the spatial kernel to locally enhance smoothness, the algorithm can adaptively adjust the width of the spatial kernel at each pixel by effectively calculating the pixel scale of the input image. If pixel i is on or near the edge, a small w_i is used. If pixel i is in a uniform or textured area, a larger w_i is used.

To adaptively adjust the size of the spatial kernel for each pixel, we used a Gaussian filter to obtain the local image information:

$$G(x) = \text{gaussian}_{y \in n} \nabla f(x - y) \tag{36.15}$$

where n represents a 3×3 window and ∇ represents the gradient. Figure 36.11 shows the gradient information of the image. The filtered output image was normalized to rescale it to the range of [0, 1].

$$m(x) = \begin{cases} \min\left(0 \le G(x) \le \frac{1}{3}\right) & \forall G(x) \in \left[0, \frac{1}{3}\right] \\ \max\left(\frac{1}{3} \le G(x) \le \frac{2}{3}\right) & \forall G(x) \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \max\left(\frac{2}{3} \le G(x) \le 1\right) & \forall G(x) \in \left[\frac{2}{3}, 1\right] \end{cases} \tag{36.16}$$

From Eq. (36.17), the specific size of the proportion w_i of the spatial kernel at pixel i can be obtained, and then the size of the window Ω_i of the spatial kernel can be obtained.

$$w_i = (w_{\max} - w_{\min}) \exp\left[-\lambda m(x)^2\right] + w_{\min} \quad \lambda > 0 \tag{36.17}$$

A Fourier-expansion acceleration algorithm was adopted to implement the scale-adaptive bilateral texture filtering and raw-image convolution. The error of the approximate values of the discrete points should be within the tolerance range. Figure 36.12 shows a comparison of the image effects on the OCC receiver using the improved center/surround retinex algorithm. Figure 36.13 shows a comparison of the characteristic grayscale-value curves for processing the imaging effects using the improved center/surround retinex algorithm.

Figure 36.12a shows the original image obtained using OCC. The overall grayscale value fluctuates greatly and the image appears dark. The local

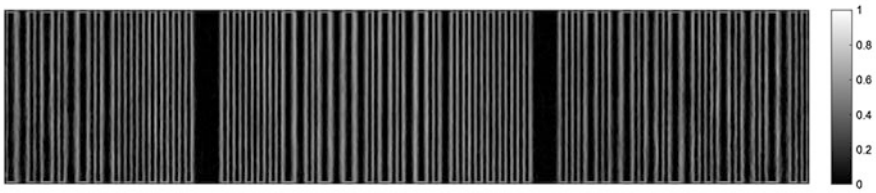


Fig. 36.11 Image gradient information

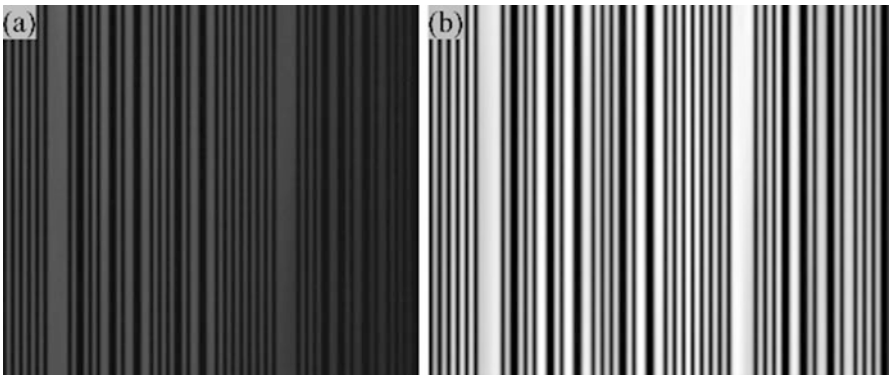


Fig. 36.12 Comparison of the image-processing effects using the improved center/surround retinex algorithm: (**a**) Before image processing and (**b**) after image-enhancement processing

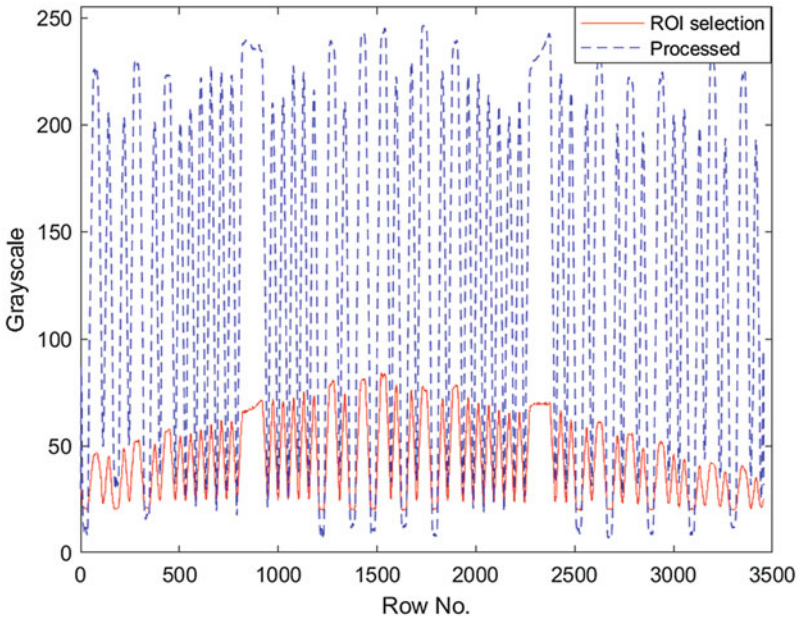


Fig. 36.13 Comparison of grayscale characteristic curves before and after image processing

grayscale values also fluctuate severely, owing to the halo effect. The blurry edges and low contrast of the bright and dark stripes seriously affect subsequent threshold decision processing.

The grayscale characteristic curve of the original image in Fig. 36.12a corresponds to the solid red curve shown in Fig. 36.13. The range of curve changes is extremely small, making it difficult to obtain accurate information using the threshold criterion. Figure 36.12b shows an image processed using the improved center/surround retinex algorithm. It is evident that the problem of fluctuations in the overall and local grayscale values of the image is resolved. The edge information of the bright and dark stripes is preserved, and the details of dark areas in the image are clearly visible.

The grayscale characteristic curve in Fig. 36.12b corresponds to the dashed blue line in Fig. 36.13. It can be seen that the distribution range and contrast of the grayscale characteristic curve processed by the algorithm in this study have significantly increased. The stripe information with smaller gray values on both sides of the image has also been improved, making subsequent decision decoding easier.

Decoding Data

A grayscale feature curve can be obtained by scanning the processed grayscale image. A simple threshold can be selected for judgment, based on the grayscale feature curve. The average of the maximum-point mean $\overline{T}_{\max}$ and minimum-point mean $\overline{T}_{\min}$ was used as the threshold for the decision to obtain a binary image, as shown in Eq. (36.18).

$$T = \frac{\overline{T}_{\max} + \overline{T}_{\min}}{2} \tag{36.18}$$

The flag-frame header should be detected to achieve synchronous reception. The bit resolution can be obtained using Eq. (36.5). The corresponding sliding window was set, based on the bit resolution, to determine the frame header. The widths of 0 and 1 of the one bit measured by the flag-frame header are W_0 and W_1, respectively, and a series of 01 is decoded according to Eq. (36.19). Then, it is decoded according to the Manchester encoding rules. Finally, the original decoded data are obtained, according to the sorting results.

$$W = \begin{cases} 0 & 0.8W_0 < W < 1.2W_0 \\ 1 & 0.8W_1 < W < 1.2W_1 \end{cases} \tag{36.19}$$

The range of W was set to $0.8W_0 < W < 1.2W_0$ and $0.8W_1 < W < 1.2W_1$ to prevent the actual measured stripe width from being poor in certain frequency bands, which could lead to a decrease in accuracy.

Experimental Results and Analysis

The bit-error rate is the most commonly used indicator of data-communication transmission quality and is influenced by various factors. As a near-field communication system, the bit-error rate of a visible-light imaging communication system is closely related to the distance and data rate of the light source.

In this part of the experiment, tilt was not characterized during the experimental phase, which means it was not a factor affecting the bit-error rate in this part of the experiment. The experiment analyzed and verified the superior performance of the improved center/surround retinex algorithm proposed in this study in eliminating global and local grayscale fluctuations, compared with other algorithms.

The experimental parameters are listed in Table 36.1. The experimental procedure is illustrated in Fig. 36.14, and the optical camera remained stable. In the experiment, the optical camera operated in the positive viewing-angle direction, and no shaking phenomena were observed. The experiment focused on two aspects:

Table 36.1 Experimental parameters

Parameter	Value			
Microcontroller	STM32F103C8T6			
Camera frame rate	30 fps			
Camera resolution	3456 × 4608			
Camera sensitivity	400			
Data cycle per bit (μs)	285	222	181	153
Data rate (kbps)	3.5	4.5	5.5	6.5
Payload bit length (bit/frame)	116	150	183	216

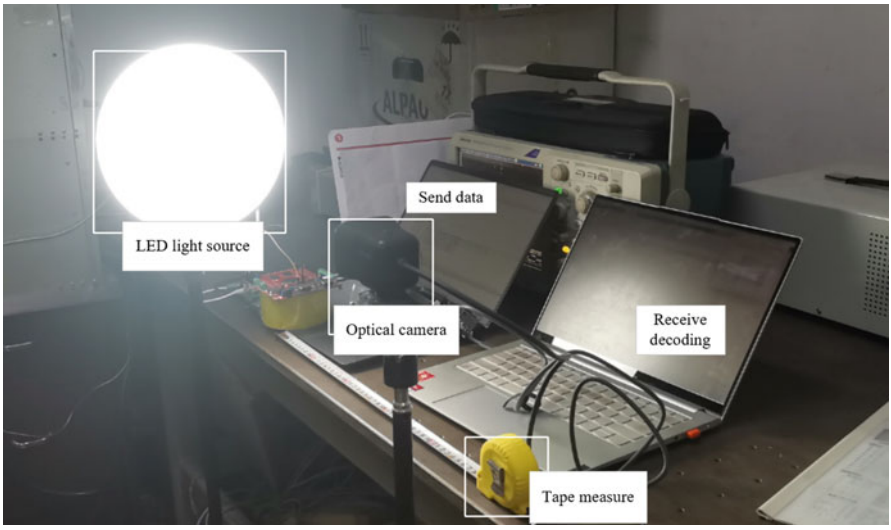


Fig. 36.14 Optical-camera communication experiment diagram

1. Different data-transmission rates at the same communication distance.
2. Different communication distances between the receiving end and the light source at the same transmission rate.

Finally, the reason why the improved center/surround retinex algorithm proposed in this study achieved good experimental results was analyzed.

Bit-Error Rate Testing

As the communication rate increases, the duty cycle of the dark stripes becomes blurred, exacerbating the extinction-ratio problem and seriously affecting the decoding. In the experiment, the transmission rate was set to 3.5–6.5 kbps, with a fixed distance of 45 cm.

The bit-error rates of the different decoding algorithms were tested, as shown in Fig. 36.15. The improved center/surround retinex algorithm proposed in this study was compared with the adaptive-contrast enhancement (ACE) algorithm [35], boundary pixels based piecewise linear function thresholding (BPBPLFT) [29],

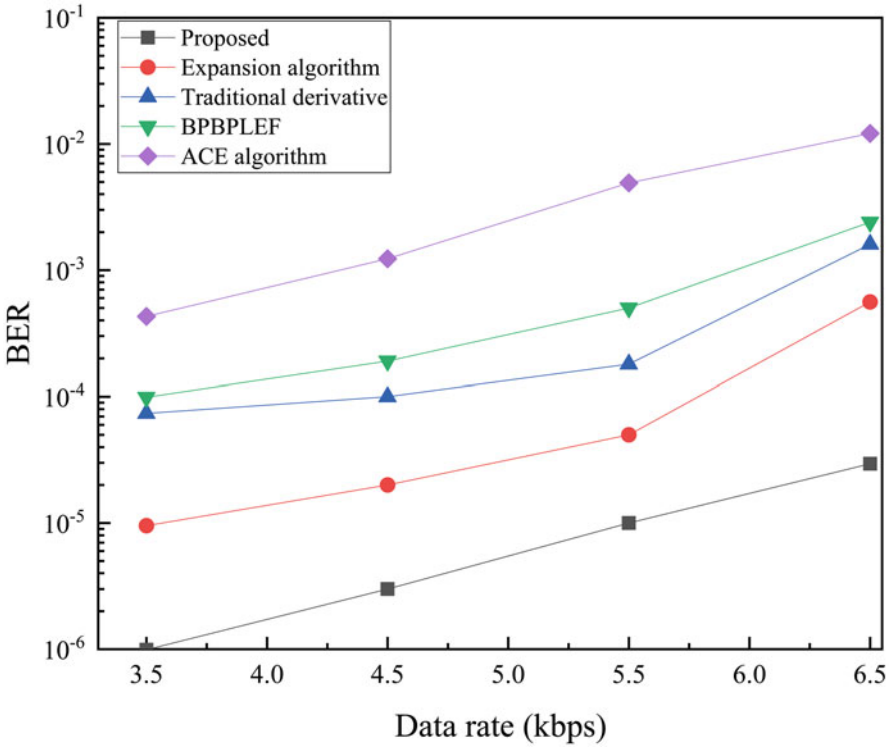


Fig. 36.15 Bit-error rates at different data rates

traditional derivative algorithm [36], and expansion algorithm [37]. The use of the adaptive-contrast enhancement algorithm and segmented linear-function threshold based on boundary pixels was limited and the traditional derivative algorithm could not obtain clear stripes when the transmission rate increased. The expansion algorithm could handle situations involving poor image quality.

As the transmission rate in the experiment gradually increased, the bit-error rate of each algorithm also increased; however, the bit-error rate of the improved center/surround retinex algorithm was always the lowest. When the communication rate was 6.5 kbps, the bit-error rate of the proposed algorithm was 2.9×10^{-5}, which is approximately one order of magnitude lower than that of the expansion algorithm.

The data rate was fixed at 5.5 kbps and the bit-error rate of each algorithm was tested within a range of 15–60 cm from the light source, as shown in Fig. 36.16. In close-range communication, as the communication distance increases, the illumination decreases, resulting in a decrease in the visible-light signal entering the CMOS sensor. Fluctuations in the overall and local grayscale values of the image received by the camera increase accordingly. The inability to obtain clear

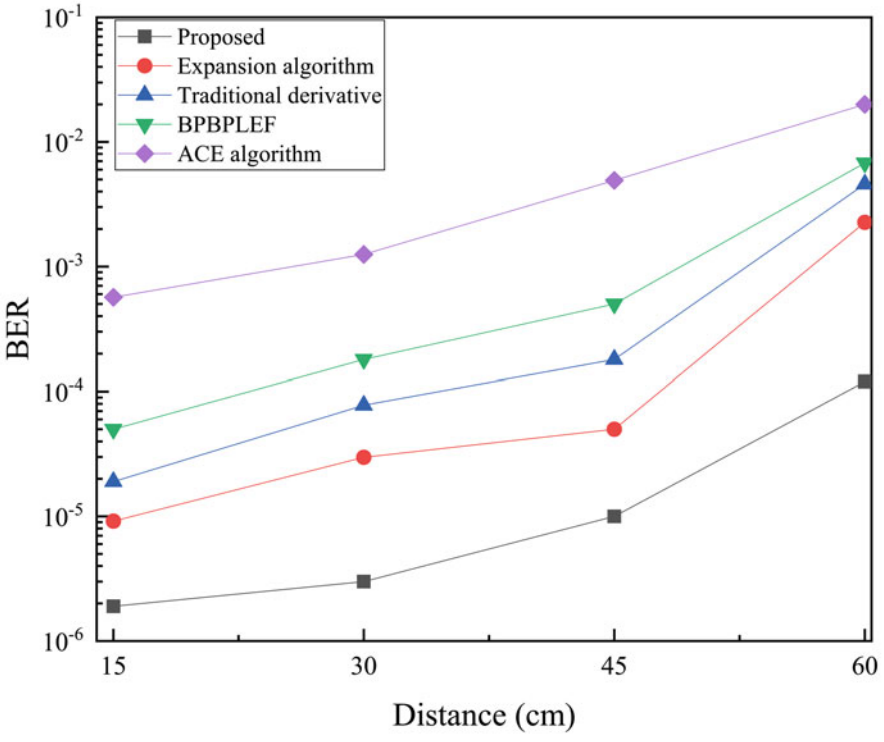


Fig. 36.16 Bit-error rates at different transmission distances

images leads to an increase in the bit-error rate, which seriously affects the decoding performance.

Based on the analysis shown in Fig. 36.16, with an increase in communication distance, the bit-error rate of each algorithm decoding increases. However, the improved center/surround retinex algorithm proposed in this study is consistently superior to the adaptive contrast-enhancement algorithm [35], segmented linear-function threshold of boundary pixels [29], traditional derivative algorithm [36], and expansion algorithm [37]. The proposed center/surround retinex algorithm has a bit-error rate of 1.2×10^{-4} at a transmission distance of 60 cm, indicating significant decoding-performance advantages.

Analysis of Experimental Results

Based on the improved center/surround retinex algorithm, the overall fluctuation of the grayscale values can be removed. Its kernel-scale adaptive bilateral texture-smoothing filter uses small spatial kernels at the edges, whereas other regions use large spatial kernels. Local smoothing removes the influence of noise, while preserving the local edge information of the image. This eliminates the problem of gray-

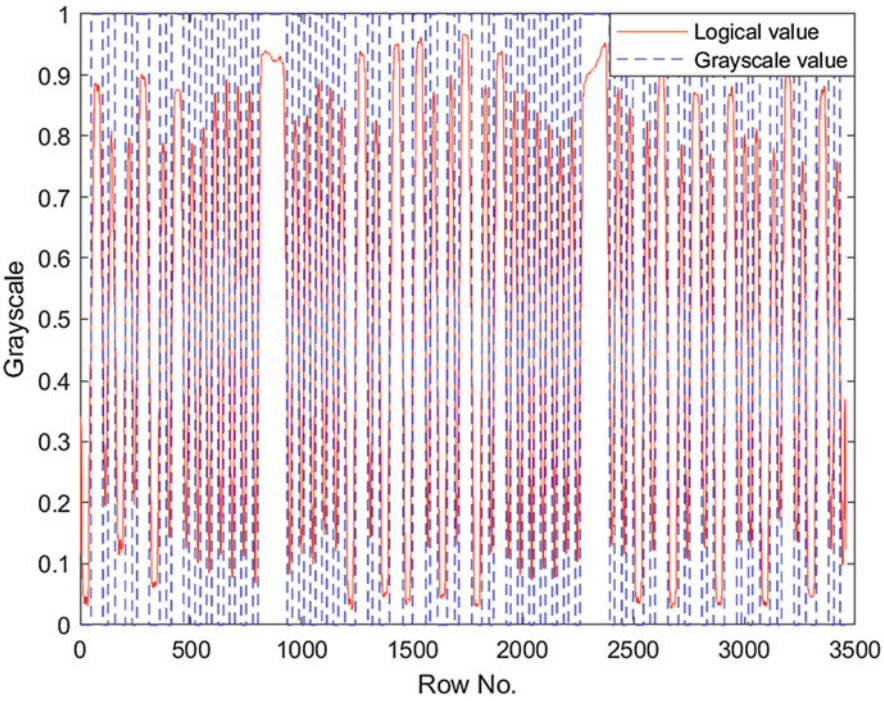


Fig. 36.17 Comparison of normalized gray values and logical values

level fluctuations caused by the halo effect and obtains an original image closest to the original transmitted data.

Subsequently, a simple threshold decision was made. The normalized grayscale values obtained using the proposed algorithm matched the corresponding logical values, and the matching effect is shown in Fig. 36.17. After obtaining a series of 01 codes, the pixel sizes of symbols 0 and 1 in the image were obtained using the frame-flag header and a decision was made. Finally, decoding and sorting produced the original transmission data.

The solid red curve in Fig. 36.17 represents the grayscale value curve processed by the improved center/surround retinex algorithm proposed in this study. The dashed blue line represents the logical value obtained from the logical-value judgment of the solid red line. It can be observed that the grayscale values correctly match the logical values. Information with smaller gray values on both sides of the image is also improved. The proposed algorithm can obtain the correct logical values to decode the data.

Conclusion

In optical-camera communication, as the system data-transmission rate increases, the number of pixel columns corresponding to the image stripes at the receiving end decreases. The traditional OCC demodulation algorithm has the disadvantage of increasing the fluctuation of the average value of grayscale pixel column with the increase of communication rate. Therefore, traditional image-processing methods produce significant errors.

At higher communication rates, the system has a higher bit-error rate, which seriously affects the communication quality. The aim of this study was to minimize the fluctuation of the grayscale values. The original data could be effectively decoded, even when the bit resolution was very small. The following conclusions were drawn from this study.

1. In response to the problems of tilt and perspective distortions, respectively caused by shaking and different receiving angles during imaging at the receiving end, corresponding correction methods were designed, based on the image characteristics, which could quickly correct the distortion problems.
2. The improved center/surround retinex algorithm simultaneously solved the problem of grayscale fluctuations in images from both global and local perspectives. The proposed algorithm eliminated the halo effect and the influence of light noise in the environment, greatly reducing the bit-error rate.
3. In the experimental verification, the improved center/surround retinex algorithm enabled the system to have a lower bit-error rate, even at higher communication rates, and was suitable for low-quality stripe images with different stripe widths, significantly improving the quality and adaptability of optical-camera communication.

References

1. Mohsan S. Optical camera communications: practical constraints applications, potential challenges, and future directions [J]. Journal of Optical Technology, 2021, 88 (12):729–741.
2. Shahjalal M, Hasan M K, Chowdhury M Z, et al. Smartphone Camera-Based Optical Wireless Communication System: Requirements and Implementation Challenges [J]. Electronics, 2019, 8(8):913–929.
3. Huang R Y, Yamazato T. A Review on Image Sensor Communication and Its Applications to Vehicles [J]. Photonics, 2023, 10(6):617–643.
4. Chow C W, Chen C Y, Chen S H. Visible light communication using mobile-phone camera with data rate higher than frame rate [J]. Optics Express, 2015, 23 (20):26080–26085.
5. Liang K, Chow C W, Liu Y. Mobile-Phone based Visible Light Communication Using Region-Grow Light Source Tracking for Unstable Light Source [J]. Optical Express, 2016, 24 (15): 17505–17510.
6. Chen C W, Chow C W, Liu Y, et al. Efficient Demodulation Scheme for Rolling Shutter Patterning of CMOS Image Sensor based VLC [J]. Optical Express, 2017, 25 (20): 24362–24367.
7. Tanaka Y, Haruyama S, Nakagawa M. Wireless optical transmissions with white colored LED for wireless home links [C]//IEEE International Symposium on Personal Indoor and Mobile Radio Communications, London, UK, 2000:1325–1329.
8. Ding D Q, Ke X Z. Visible Light Communication and Research on Its Key Techniques [J]. Semiconductor Optoelectronics, 2006, 27(2):114–117.
9. Nguyen T, Islam A, Yamazato T, et al. Technical issues on IEEE 802.15.7m image sensor communication standardization [J]. IEEE Commun. 2018, 56(2):213–218.
10. Danakis C, Afgani M, Povey G, et al. Using a CMOS camera sensor for visible light communication [C]//IEEE Globecom Workshops, Anaheim, CA, USA, 2012.
11. Takai I, Ito S, Yasutomi K, et al. LED and CMOS Image Sensor Based Optical Wireless Communication System for Automotive Applications [J]. IEEE Photonics Journal, 2013, 5(5):6801418.
12. Chen S H, Chow C W. Color-filter-free spatial visible light communication using RGB-LED and mobile-phone camera [J]. Optics express, 2014, 22 (25):30713–30718.
13. Chow C W, Chen C Y, Chen S H. Visible light communication using mobile-phone camera with data rate higher than frame rate [J]. Optics Express, 2015, 23 (20):26080–26085.
14. Chow C W, Chen C Y, Chen S H. Enhancement of signal performance in LED visible light communications using mobile phone camera [J]. IEEE Photonics Journal, 2015, 7(5):1–7.
15. Liu Y, Chow C W, Liang K, et al. Comparison of thresholding schemes for visible light communication using mobile-phone image sensor [J]. Optics Express, 2016, 24(3):1973–1978.
16. Liu Y, Liang K, Chen H Y, et al. Light Encryption Scheme Using Light-Emitting Diode and Camera Image Sensor [J]. IEEE Photonics Journal, 2016, 8(1):1–7.
17. Liu Y, Chen H Y, Liang K, et al. Visible light communication using receivers of camera image sensor and solar cell [J], IEEE Photon. J., 2016, 8(1).
18. Liang K, Chow C W, Liu Y. Mobile-phone based visible light communication using region-grow light source tracking for unstable light source [J] Optics Express, 2016, 24 (15): 17505–17510.
19. Liang K, Chow C W, Liu Y, et al. Thresholding schemes for visible light communications with CMOS camera using entropy-based algorithms [J]. Optics Express, 2016, 24(22): 25641–25646.
20. Li Z Z, Zhang Z S, Yuan Q Z, et al. Digital image processing in led visible light communications using mobile phone camera [C]//IEEE International Conference on Network Infrastructure and Digital Content, Beijing, China, 2016:239–243.
21. Dai Y H, Tang X. Optical camera communication using high-resolution image sensor within a single frame [J]. Journal Of Nanjing University Of Information Science & Technology (Natural Science Edition), 2017, 9(02):185–191.

22. Chen C W, Chow C W, Liu Y, et al. Efficient demodulation scheme for rolling-shutter-patterning of CMOS image sensor based visible light communications [J]. Optics Express, 2017, 25(20):24362–24347.
23. Wang W C, Chow C W, Wei L Y, et al. Long distance non-line-of-sight (NLOS) visible light signal detection based on rolling-shutter-patterning of mobile-phone camera [J]. Optics Express, 2017, 25(9):10103–10108.
24. Zhang Z S, Zhang T T, Zhou J, et al. Performance Enhancement Scheme for Mobile-Phone Based VLC Using Moving Exponent Average Algorithm [J]. IEEE Photonics Journal, 2017, 9(2):1943–0655.
25. Lee J W, Kim S J, Han S K. Adaptive window thresholding for noise-robust photo detection in OCC [J]. Optics Communications, 2018, 426(1):623–628.
26. Shi J, He J, Jiang Z W, et al. Enabling user mobility for optical camera communication using mobile phone [J]. Optics Express, 2018, 26 (17):21762–21767.
27. Guan W P, Wu Y X, Xie C Y, et al. Performance analysis and enhancement for visible light communication using CMOS sensors [J]. Optics Communications, 2018, 410:531–551.
28. Sun H W. Design and Application of Visible Light Communication System Using Smartphone [D]. Dalian: Dalian Maritime University, 2018.
29. Zhang Z S, Zhang T T, Zhou J, et al. Thresholding Scheme Based on Boundary Pixels of Stripes for Visible Light Communication With Mobile-Phone Camera [J]. IEEE Access, 2018, 6: 53053–53061.
30. Zhang Z S, Qiao Y J, Zhang T T, et al. Fractional Weight Moving Average Based Thresholding Scheme for VLC with Mobile-Phone Camera [J]. IEEE Photonics Journal, 2019, 11(1):1–8.
31. Li T Y, Chi X F, Shi H Y, et al. Rolling shutter aided optical camera communications with increasing communication distance [J]. Optoelectronics Letters, 2019, 15(5):363–367.
32. Zhang L, Wang X D, Wu N. Realization scheme of optical camera communication system based on RGB-LED [J]. Optical Communication Technology, 2019, 43(5):26–31.
33. Zhong S H, Zhu Y H, Chi X F, et al. Optical Lensless-Camera Communications Aided by Neural Network [J]. Applied Sciences, 2019, 9(16):3238–3252.
34. Zhu Y K, Li C L, Shan J X, et al. Research On Debugging Information Secure Transmission For Wireless Sensor Nodes Based On Optical Camera Communication [J]. Information Technology and Network Security, 2019, 38 (09):14–17.
35. Liu X J, Feng L F, Chen D Y. An efficient demodulation scheme for VLC signal image processing [J]. Optoelectronics Letters, 2020, 16(3):211–215.
36. Meng Y M, Chen X B, Pan T H, et al. HyperSight: A Precise Decoding Algorithm for VLC With Mobile-Phone Camera [J]. IEEE Photonics Journal, 2020, 12(4):1–11.
37. Li P P, Feng L F, Huang S Y. New Type of Optical Camera Communication Signal Demodulation Scheme [J]. Acta Optica Sinica, 2021, 41(23):68–74.
38. Bai B, Su B, Li J L, et al. Data Frame Structure for Optical Camera Communication System [J]. Acta Electronica Sinica, 2021, 49 (10):1873–1880.
39. Wu J, Chi X F, Ji F L, et al. Multi-level modulation scheme based on PAK algorithm for optical camera communications [J]. Optoelectronics Letters, 2022, 18(3):152–157.
40. Dong K, Ke X Z, Wang M J. Equalization of camera-based channel to mitigate uncertain sampling for optical camera communications [J]. Optics Express, 2022, 30 (26): 47776–47791.
41. Zhang Y P, Zhu D Y, Ma J M, et al. Subway Train Positioning Based on Monocular Vision and Optical Camera Communication [J]. Laser & Optoelectronics Progress, 2022, 59 (16):69–78.
42. Chen Y, Han Z Z, Liu H L, et al. Optical Camera Communication Method Based on Reference Fringe Width Judgement [J]. Journal of Electronics & Information Technology, 2022, 44 (08): 2629–2638.
43. Jiang K Y, Chi X F, Ji F L, et al. Research on mobile phone swaying and receiving position in optical camera communication [J]. Optoelectronics Letters, 2022, 18(2):86–90.
44. Zhang Y P, Zhu X Q, Zhu D Y, et al. Train Positioning Using Optical Camera Communication with BP Neural Network [J]. Chinese Journal of Lasers, 2023, 50(5):124–134.

45. Dong K, Ke X Z, Zhang X Z, et al. Parametric modeling and experimental measurement of rolling shutter characteristics for optical camera communication using under sampled modulation [J]. Applied Optics, 2022, 61(27):7838–7845.
46. He J, Jiang Z W, Shi J, et al. A novel column matrix selection scheme for VLC system with mobile Phone camera [J]. IEEE Photonics Technology Letters, 2019, 31(2):149–152.
47. Ansari I, Shim J. Vehicle Manufacturer Recognition using Deep Learning and Perspective Transformation [J]. Journal of Multimedia Information System, 2019, 6(4):235–238.
48. Mukhopadhyay P, Chaudhuri B B. A survey of Hough Transform [J]. Pattern Recognition, 2015, 48(3):993–1010.
49. Oishi S, Fukushima N. Retinex-Based Relighting for Night Photography [J]. Applied Sciences, 2023, 13(3):1719–1737.
50. Ruhela R, Gupta B, Singh L S. An efficient approach for texture smoothing by adaptive joint bilateral filtering [J]. The Visual Computer, 2023, 39:2035–2049.

Indoor Visible Light Positioning Technology

37

Contents

Introduction 1414
Visible Light Positioning Systems 1414
Geometric Ranging 1414
Scene Analysis 1416
Proximity Sensing 1417
Image Sensor Imaging 1417
Domestic and Foreign Research Progress on Indoor Visible Light Localization 1419
Foreign Research Progress 1419
Domestic Research Status 1422
Research Progress on Key Technologies 1424
LED Layout 1424
Light Source Modulation Modes 1426
Intelligent Optimization Algorithms 1427
Multifeature Fusion Localization 1427
Summary 1428
References 1430

Abstract

Visible light positioning (VLP) uses LED lamps for lighting to achieve indoor high-precision positioning, which has the preferences of cost-effective, no electromagnetic intermeddling, high security, and convenient deployment. In this paper, firstly, the indoor positioning system and routine visible light positioning algorithms are retrospected, and the current development status is acquainted. Then, some pivotal technologies to embellish the performance of the overall positioning system are discussed, which provides a footing for further research on the white LED indoor location system. Finally, on account of the subsisting research results, the future research direction of indoor visible light positioning technology is put forward.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_37

Keywords

Visible light communication · Visible light positioning · Indoor positioning

Introduction

At present, indoor locations mostly adopt Bluetooth, infrared, ultrasonic, radio-frequency (RF), etc., for wireless communication. Most wireless location technology is affected by the sheltering and multipath effects of buildings, and the positioning accuracy is very low in indoor environments. Thus, it cannot meet the demands of indoor location services.

Visible light light-emitting diode (LED) indoor location technology is superior. It only requires the transformation of lamps and lanterns and the addition of chips, to both light an area and provide location services. Visible light positioning technology has the advantages of low cost, green environment protection, and strong resistance to electromagnetic interference. It is very suitable for complex indoor environments, especially for obtaining accurate positions in places such as mines, hospitals, and nuclear power plants [1].

In this chapter, the key technologies to enhance the performance of indoor positioning systems are discussed because the positioning systems used in different application scenarios must meet different standard requirements, such as cost and reliability.

Visible Light Positioning Systems

A visible light LED-based indoor positioning system has the same system architecture as other positioning technologies, such as WiFi-based or Bluetooth-based systems. A typical prototype of the system is shown in Fig. 37.1.

A visible light indoor positioning system consists of two parts: one or more transmitters and one receiver [2]. Depending on the terminal location, they are divided into a photodiode (PD)-based non-imaging location system and an image sensor (IS)-based imaging location system. The system block diagram is shown in Fig. 37.2.

According to the different location methods, visible light location algorithms can be divided into geometric ranging, scene analysis (fingerprint location), proximity sensing, and image sensor imaging methods.

Geometric Ranging

By measuring the distance between the reference point and the target to be measured, the position coordinates of the object to be measured can be estimated using the geometric relation between the reference point and the target. It includes four

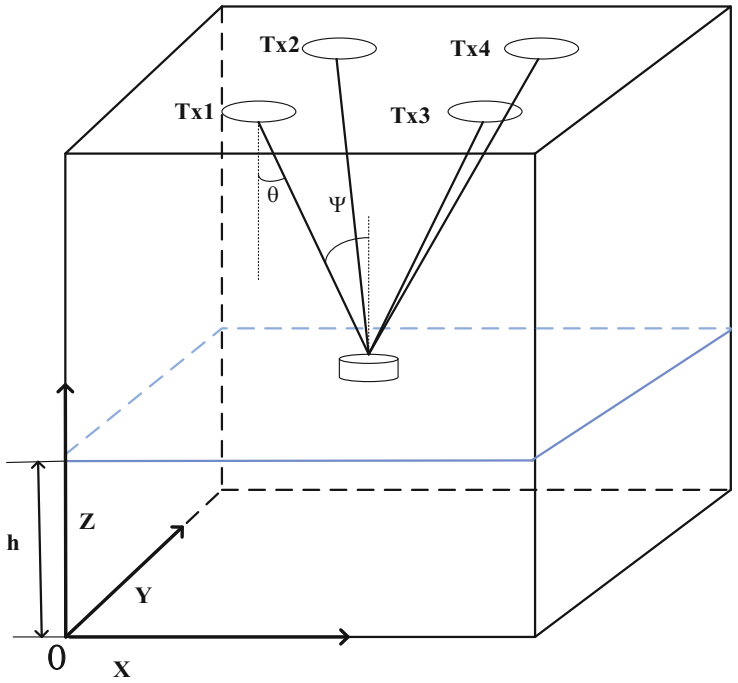


Fig. 37.1 Indoor positioning system based on LEDs

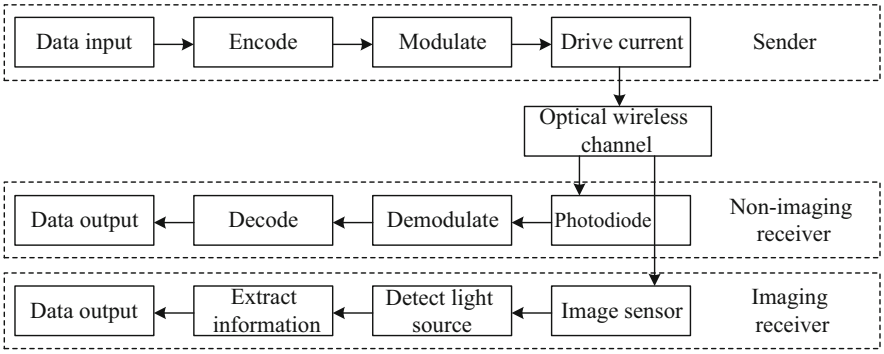


Fig. 37.2 Location system block diagram

methods: time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA), and received signal strength (RSS).

The TOA estimates the distance between the LED and the target, based on the signal arrival time. For visible light, the distance between the transmitter and receiver is calculated by multiplying the signal propagation delay by the speed of light [3]. The TDOA-based positioning system estimates the position of the receiver based on

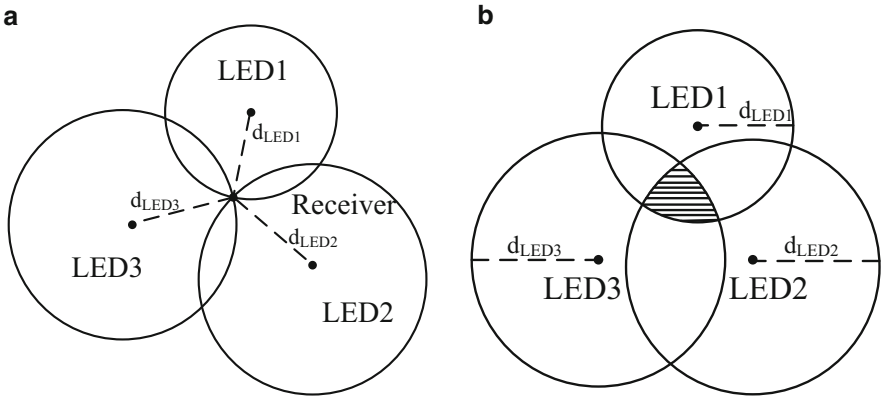


Fig. 37.3 Intersection of three-sided measurements. (**a**) Triangulation using only one intersection. (**b**) Triangulation intersecting one surface

the time difference of the arrival of signals from multiple LEDs and only requires time synchronization between the LEDs.

AOA uses known LED coordinates and the angle between the target and LED to determine the position, which requires at least two LED lights. The positioning accuracy is high for indoor locations. This method eliminates time synchronization without path loss or interference caused by the background or reflection; however, measuring the angle requires additional hardware, which is an important drawback.

RSS-based distancing typically uses direct detection in the receiver while measuring at least three known transmitter position coordinates to determine the unknown receiver location. In general, RSS follows the Lambert model. The receiver measures the strength of the transmitted signal, calculates the distance to each transmitter, and draws the possible location region based on the radius. The position is at the intersection of the three circles, as shown in Fig. 37.3a.

Owing to quantization and measurement errors, in most cases, the circles do not intersect perfectly, and the points of intersection are at the edges of the overlapping areas, as shown by the shadows in Fig. 37.3b. This intersection with a plane requires at least four LED emitters to determine unknown position coordinates [4].

Scene Analysis

Scene analysis, also known as fingerprint identification, is a non-ranging location algorithm that includes online and offline stages, as shown in Fig. 37.4. In the offline-investigation stage, specific information about the indoor environment is collected and processed. The fingerprint features include the light intensity [5], pulse response time [6], and extinction ratio [7].

In the online stage, the location of the unlocated point is estimated using a matching model, which includes three main algorithms: deterministic localization

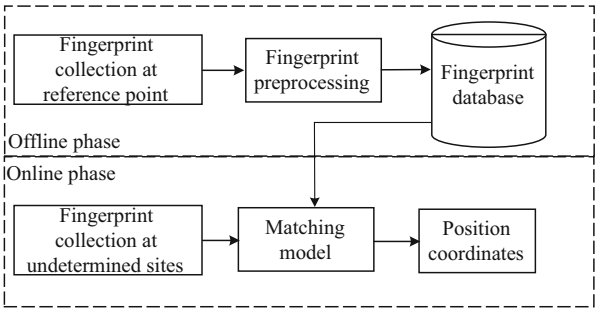


Fig. 37.4 Principle of fingerprint localization [12]

(nearest neighbor [8]), probabilistic localization (naive Bayes [9]), and neural network localization (backpropagation (BP) neural network [10]). In addition, correlation is a new fingerprint-matching technique proposed in the literature [11]; when a mobile device is closer to a specific base station, the correlation between the mix of received signals from all base stations and the transmitted signal from that base station is higher.

The key to the fingerprint location method is to establish a fingerprint database. The location accuracy depends on the size of the area and the degree of feature identification; however, the portability of the system is an unavoidable drawback.

Proximity Sensing

The proximity sensing method, also known as the LED method, is a rough localization method, as shown in Fig. 37.5. Each LED lamp is controlled by a terminal that emits light with unique characteristics. By comparing the light intensity of the specific signal at the receiver, the approximate position of the target can be determined according to the signal intensity [13]. This method has a simple system structure and is easy to realize, in terms of coding and modulation. This is a rough location, and its location precision is closely related to the distribution density of the transmitter.

Image Sensor Imaging

The image sensor imaging method is based on the theory of aperture imaging, with the help of image sensor parameters, and object and image point coordinates of the relative position relationship, to obtain the image sensor position information. The current receiver can use single or multiple image sensors. As shown in Fig. 37.6, the monocular camera image positioning system uses a single camera to capture images of multiple LEDs at the receiving end of the system. The position of the image sensor is determined by identifying the positions of multiple LEDs.

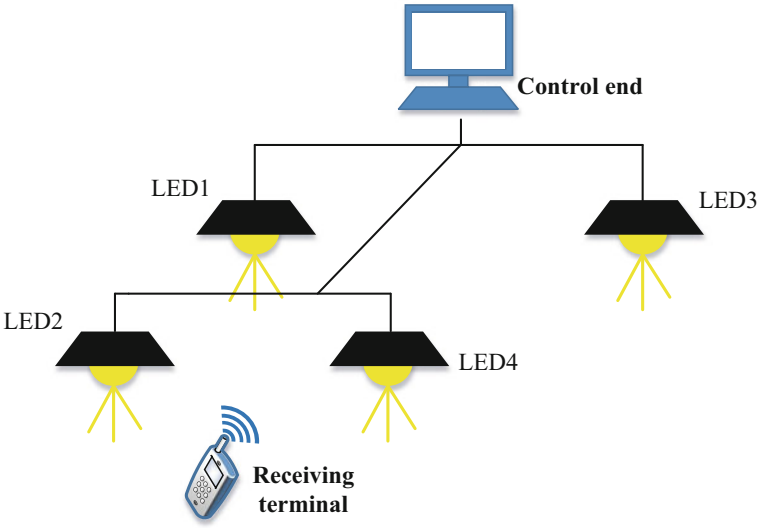
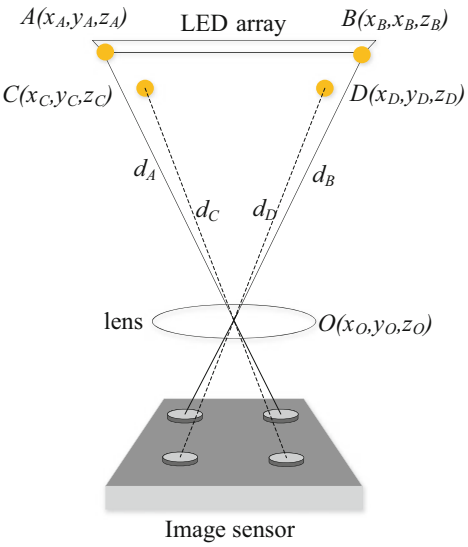


Fig. 37.5 Schematic of the LED positioning method [14]

Fig. 37.6 Location system using a monocular camera



Owing to the angle error caused by a single image sensor, some researchers use two image sensors, such as in Fig. 37.7, to determine the position, effectively reducing the angle error. The positioning accuracy is higher and more flexible; however, this method requires complex equipment and has a high cost. Some

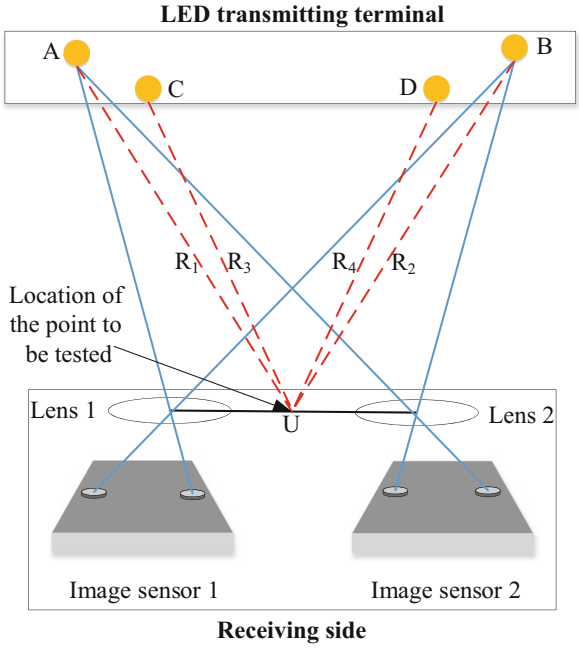


Fig. 37.7 Location system using a binocular camera [15]

researchers also use inertial sensors and other auxiliary hardware to reduce the angle error of a single image sensor.

The above discussion is for 2D positioning. It is assumed that all LEDs are installed at the same height and that the z coordinates of the receiver are known. It should be noted that when the installation error exceeds 5%, the positioning accuracy is affected; there is an error of 20 cm higher in [16]. The principles of the visible light localization algorithms are briefly described above, and the performances of these algorithms are compared and summarized in Table 37.1.

Domestic and Foreign Research Progress on Indoor Visible Light Localization

Foreign Research Progress

Since the introduction of the E-911 positioning standard by the Communications Commission in 1996, indoor positioning technology has rapidly developed [18]. In 1999, Pang [19] first proposed the theory of visible light communication (VLC) technology based on light-emitting diodes (LEDs). In 2007, the Japanese Electronics and Information Technology Industries Association (JEITA) developed the CP-1221 (VLC system) and CP-1222 (visible light ID system) standards; CP-1222 is the standard in the field of visible light positioning (VLP) [20].

Table 37.1 Performance comparison of visible light localization algorithms [17]

Method	Precision	Coverage	Cost	Multipath interference	Extra hardware	Advantage	Disadvantage
LED	Depends on the light spacing	Good	Low	Yes	No	Low cost	Low accuracy
AOA	Medium	Good	High	Yes	Yes	No bandwidth required	(1) High cost; (2) complex; (3) requires antenna arrays
TOA/ TDOA	High	High	High	Yes	Yes	High precision	(1) High cost; (2) requires precise time synchronization
RSS	Medium	Good	Medium	Yes	No	Low cost	(1) Low precision; (2) susceptible to multipath fading
Fingerprinting	Depends on the grid-partition density	Good	Medium	No	No	Low cost	(1) Requires extensive calibration; (2) map maintenance is complex
Image method	High	Good	High	No	Yes	Strong anti-interference	(1) Complex algorithm; (2) high cost

In 2010, when researchers began focusing on visible light-based localization techniques, Sertthin used a single photodiode (PD) for localization, experimented with different field-of-view (FOV) angles at the receiver, and obtained an inverse relation between the precision and FOV [21]. In 2012, Prince et al. implemented a hybrid system using the RSS and AOA algorithms that used several PDs with different FOVs between 70° and 100°. The results showed that the optimal angle was 90°, and the positioning error was 13.5 cm [22]. Thus, in VLP system design, the choice of FOV needs to be weighed against the specific model. The same is true with respect to the image sensor resolution, as shown in Figs. 37.8 and 37.9.

In 2011, Giulio Cossu implemented short-distance communication using orthogonal frequency division multiplexing (OFDM) technology, combined with a triangulation algorithm, to achieve a communication rate of 300 Mbit/s at a close range, and the feasibility of localization was presented [23]. Information systems (IS) combined with smartphones to move VLC location technology from the lab to the commercial market.

In 2013, South Korea's Renaissance Laboratory developed emart, an app for supermarkets, that provided in-store navigation for consumers. Panasonic has 200 optical transmitters on the third floor; visitors wear special receivers in their ears and connect to their phones via Bluetooth to reach their destinations, according to the arrows provided by the location app. In 2013, Sekercioglu analyzed the accuracy of VLC using the arrival time algorithm. The theoretical limit of the estimation accuracy, calculated using the Cramer–Rao lower bound (CRB), showed that the error range was 2–6 cm [24].

In complex environments, multipath reflection is an important factor affecting the accuracy of indoor positioning. In 2016, Aminikashani et al. used a deterministic and

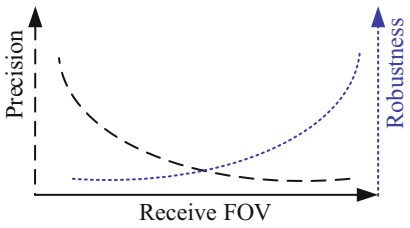


Fig. 37.8 Tradeoffs with respect to the FOV at the receiving end [11]

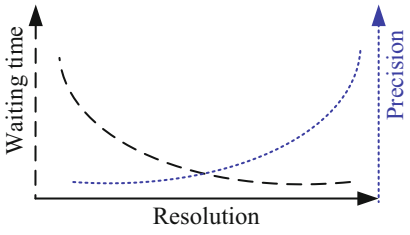


Fig. 37.9 Tradeoffs with respect to the image sensor resolution [11]

improved Monte Carlo method to calculate the impulse response to the multipath reflection effect. The accuracy of the system reached 0.558 m [25], using the calibration method to reduce the positioning error. However, the combination of very high reflectivity and thermal noise with shot noise may result in the highest signal being received as a diffuse signal in some locations. In 2019, Tran de-noised the signal and eliminated the low-intensity reflective signal using machine-learning (ML) classification and regression to improve the VLP, improving the positioning accuracy by 52.55% [26].

With the development of big-data processing technology, intelligent optimization algorithms have been widely used. In 2018, Hsu et al. first applied a neural network in a localization experiment in a visible light chamber and demonstrated continuous target tracking. The distance between the LED lamp and receiver was 2.5 m, and the average position error reached 3.65 cm [27].

In 2020, Yu-Chun proposed and demonstrated a VLP system using sigmoid data preprocessing and applied it to two regression-based ML algorithms with an average experimental positioning error of 2 cm [28]. The optimization algorithm improved the indoor location accuracy and the data-utilization rate. In 2023, Ahmed's proposed location system verified that support vector regression (SVR) was 46% better than a one-neuron neural network and 0.65% better than a 10-neuron neural network in terms of location error. They determined that it was appropriate to integrate it into mobile devices and apply it to real-life problems [29] (Table 37.2).

Domestic Research Status

In 2006, Ding et al. from Xi'an University of Technology first proposed the use of OFDM to reduce intersymbol interference (ISI) in indoor location systems based on VLC. In 2007, Koh et al. began to study LED light source layouts with minimal light-power fluctuations [30], providing a basis for visible light positioning. The LED luminescence model was important for VLP design. In 2010, a mathematical luminescence model of a general white LED was proposed. The similarity coefficient between the model and manufacturer data was more than 95%, and the model precision was high [31].

In 2014, Zhang theorized the CRB expression of the received signal strength index (RSSI) algorithm for evaluating the positioning error performance and considered the effect of multipath reflection on the positioning accuracy [32]. In 2015, Wu et al. proposed an improved RSS-based triangulation algorithm that achieved very low error fluctuations with or without occlusion, with a positioning accuracy of less than 3.5 cm [33].

In 2016, Guan applied code-division multiple access (CDMA) to indoor visible light communication, using the orthogonality of the spread-spectrum code and diversity receiver technology. The algorithm verified that more than 88% of the system positioning accuracy was within 5 cm [34]. In 2018, Ye et al. proposed the use of wavelet denoising and a cross-correlation function detection method to extract useful signals at the receiving end because of the large RSS error in measuring the

Table 37.2 Foreign research status of indoor visible light localization

Year	Researcher/ Organization	Research progress	Precision	Receiving end
1999	Grantham Pang	The theory of visible light communication based on LEDs was presented for the first time		
2007	JEITA	The CP-1221 and CP-1222 standards were established		
2010	Sertthin	The relationship between accuracy and FOV was obtained		PD
2011	Giulio Cossu	The feasibility of short-distance communication and location was proposed		PD
2012	Prince	The RSS and AOA positioning systems found the best FOV value	13.5 cm	PD
2013	Sekercioglu	Analyzed the TOA algorithm using the CRB	2–6 cm	PD
2016	Aminikashani	The Monte Carlo method calculated and calibrated the impulse response	0.558 m	PD
2018	Hsu	Neural networks first applied to VLP and tracking is realized		PD
2019	Tran	Used ML classification and regression to improve VLP	Positioning accuracy improved by 52.55%	PD
2020	Chow	Developed a sigmoid function data preprocessing VLP system	2 cm	PD
2023	Ahmed	Machine-learning approach is used in conjunction with RSS	The SVR location error was 46% better than that of a neural network	PD

distance between the receiving end and the sending end. The positioning error could be stabilized within 1 cm, and the accuracy was 10 times higher than that without treatment [35].

At present, the dynamic-location problem is seldom involved in imaging locations based on VLC. In 2017, Fang achieved an average positioning time of 22 ms by designing a lightweight image-processing algorithm that supported 5 Mbit/s real-time navigation with both positioning accuracy and real-time performance [36].

One of the biggest problems when using a camera is lens distortion, which can occur with any lens, as shown in Fig. 37.10. In 2018, Wang used camera calibration, monocular vision, and location technology to partition an indoor environment and applied a calibration process to solve the distortion problem. They achieved an average point-of-view location time of less than 5 s and a positioning accuracy of

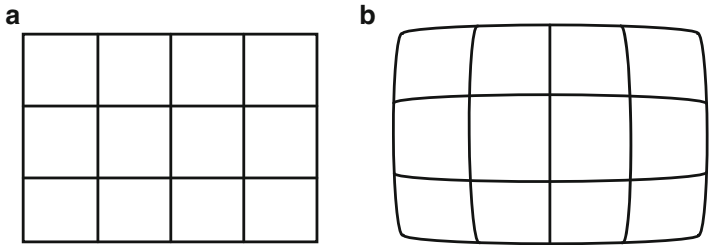


Fig. 37.10 Effect of lens distortion on an image. (**a**) Original image (**b**) Distorted image

less than 0.5 m [37]. Wang combined the advantages and disadvantages of photodetectors and camera receivers, using the photodetector to receive ID information and the camera to receive visual information; it achieved 3D positioning at the decimeter level [38] (Table 37.3).

At present, most studies on VLP systems only focus on positioning accuracy; however, robustness and real-time performance are often neglected. In 2019, Chen proposed an improved extreme learning machine (ELM) algorithm based on fingerprints, which had fast-localization and fast-learning abilities. In 2020, Chen introduced the K-means++ clustering algorithm to cluster fingerprints with similar signal strengths into one class and then used a weighted k-nearest neighbor (WKNN) algorithm to calculate the location error, to solve the problem of large errors in most indoor locations.

In 2021, Zhang proposed a clustering fingerprint location algorithm to screen out the edge area and enlarge it effectively; the edge-region error was reduced [40]. In 2022, Qin proposed to use two fingerprint databases to train two Elman neural networks to calibrate predictions. An average positioning error of 7.13 cm and average positioning time of 0.24 s were obtained [41].

Research Progress on Key Technologies

Indoor visible light localization is easy to realize; however, it is easily influenced by the outside environment, which leads to large positioning errors, including multipath reflection, inter-cell interference, and the presence of ambient light. Many researchers have focused on developing VLP systems with high precision, reliability, and reasonable costs. Next, we discuss the four main VLP technologies: LED layout, modulation, intelligent optimization algorithms, and multifeature fusion location.

LED Layout

Compared with other indoor wireless location technologies, indoor locations based on visible light must include reliable lighting to obtain an accurate location. The LED light source layout is the basis of information transmission. The LED layout

Table 37.3 Domestic research status of indoor visible light localization

Year	Researcher/ Organization	Research progress	Precision	Receiving end
2006	Ding Deqiang	VLC uses OFDM to reduce intersymbol interference (ISI)		
2014	Zhang	The CRB expression of the RSSI algorithm is derived theoretically		
2015	Wu Nan	The trilateral location method is improved, considering the LED topology	3.5 cm	PD
2016	Guan Weipeng	Applied CDMA to indoor VLC	5 cm	PD
2017	Fang Junbin	Designed a lightweight image-processing algorithm	5 Mbit/s real-time navigation	IS
2018	Wang Yong	Partitioned and applied calibration to solve the distortion problem	0.5 m	IS
2018	Wang	The PD received the ID information, and the camera received the visual information	Minutes and seconds	PD and IS
2018	Wei Yezi	Wavelet denoising and cross-correlation function detection used to extract useful signals	1 cm	PD
2019	Chen	Proposed an improved ELM algorithm based on fingerprints	0.07 m	PD
2020	Chen	Introduced K-means++ clustering algorithm		PD
2021	Zhang Rui	Filtered out and enlarged the edge area	Overall accuracy improved by 9.4%	PD
2022	Qin Ling	Trained two Elman neural networks using a fingerprint database	7.13 cm	PD

can also affect the positioning error, and the average positioning error can be reduced by increasing the LED coverage area.

In most cases, a single LED light source cannot meet the international requirements of 300 lx–1500 lx indoor illumination. Reference [42] states that the total number of LED emitters and PDs must be greater than three. If the number of used LED transmitters is greater than the number of PD receivers, a better positioning performance can be obtained. Multi-LED light sources are widely used in indoor lighting; hence, the illumination uniformity, power consumption, and communication reliability must be considered when designing LED arrays.

In the past, a large number of studies only considered the influence of a single layout model on the overall accuracy. They did not fully consider the relationship between the layout model and positioning accuracy. The signal strength of the receiver fluctuates significantly, which affects the positioning accuracy. Ding et al.

of Xi'an University of Technology [1] determined the receiving surface layout with the least optical power fluctuation in a $5 \times 5 \times 3$ m^3 room configuration. The average value was 1.5 dBm and the variance was a minimum of 0.02 dBm. Most current VLP experiments use this layout.

To meet the basic indoor illumination requirements, Shen et al. [43] determined the LED positions corresponding to the minimum positioning error of the CRB on the surface to be measured. Zhao [44] designed the $4 + 1$ LED illumination-compensation distribution mode and constructed an optimization function considering the illuminance mean-square deviation and the average bit-error rate. They obtained an illuminance uniformity of 93.23% under the optimal point and an average bit-error rate of 6.45×10^{-7}.

Guan used a new algorithm to verify that the location accuracy of an isosceles triangle was better than that of a right triangle, because the LEDs were more evenly distributed. Guan [45] compared the precision-stability interval and limit mean error of different models for the basic and dense layout models of a multilight source in a VLP system. However, it was not significant, which meant that the stability of the positioning system could be improved using an appropriate LED layout.

Compared with multiple LED layouts, LED-array layouts often achieve better positioning performance, and uniform and symmetrical light sources can provide a more stable illumination environment; however, the improvement of the positioning performance of the entire system is limited.

Light Source Modulation Modes

Light-emitting diodes are used for lighting and communication purposes; therefore, appropriate light source modulation should be used to modulate the signal containing the information to switch LEDs at the desired frequency. This means that VLC modulation methods must not only achieve data rates but also meet lighting requirements [46]. The modulation modes commonly used in visible light communication include on–off keying (OOK) [47], pulse-position modulation (PPM) [48], OFDM [49], carrierless amplitude and phase (CAP) modulation, and color-shift keying (CSK) [50]. OOK and PPM have simple structures and are easy to implement, whereas OFDM has better performance in improving spectrum utilization and overcoming the multipath effect.

In 2004, Komine et al. used the OOK scheme, but two important issues were not considered: dimming control and flicker control. On this basis, some studies combined OOK with the Manchester code, which effectively avoids continuous signals and affects normal lighting. An indoor positioning system based on PD requires multiple light sources to transmit position information; the transmitting signals overlap in the time and frequency domains, which affects the positioning performance. Therefore, the advantages of multiplexing in non-imaging VLP systems have been demonstrated.

One paper [51] proposed to use asymmetric limiting optical OFDM (ACO-OFDM), DC biasing optical OFDM (DCO-OFDM), and asymmetric limiting

Table 37.4 Performance comparison of different modulation modes

Modulation mode	Rate	Dimmer	Flicker problem	Anti-interference	Complexity
OOK	Low	Supported	High	Low	Low
PPM	Medium	Supported	Low	Medium	Medium
OFDM	High	Not supported	Low	High	High
FSK	Medium	Not supported	Low	Medium	Lower
CAP	Medium	Supported	Low	High	High
CSK	Medium	Supported	Low	High	Low

DC biasing optical OFDM (ADO-OFDM) to solve the localization error caused by multipath effects. In 2015, Xiang Yang proposed a visible light localization scheme based on an ACO-OFDM pilot signal. The positioning accuracy was improved, compared with OOK [52]. A comparison of the performances of different modulation modes is presented in Table 37.4.

Intelligent Optimization Algorithms

Research has shown that combining VLP and intelligent optimization algorithms can improve the location performance in 2D and 3D locations. Reference [53] divides the maximum-likelihood technique into four types: supervised learning, unsupervised learning, reinforcement learning, and deep learning, as shown in Fig. 37.11.

The supervised learning model with classification and regression functions is very suitable for improving positioning accuracy and reducing the computing time to meet the real-time requirements of practical applications. In the supervised learning model, common algorithms include artificial neural networks (ANNs), k-nearest neighbor (KNN), decision tree (DT), random forest (RF), and support vector regression (SVR). It has been proven that the supervised learning method is more accurate and stable than the traditional method.

The intelligent algorithms used in localization are summarized in Table 37.5. The algorithm optimization process needs to be further explored.

Multifeature Fusion Localization

In an indoor environment, a single positioning technology cannot cover all the areas, and the positioning accuracy is limited. In addition to parameter estimation using only a single attribute of the received VLC signal (e.g., TOA, RSS, or AOA), there are hybrid schemes that combine several signal attributes to estimate location-dependent parameters. Compared with traditional techniques, this hybrid approach

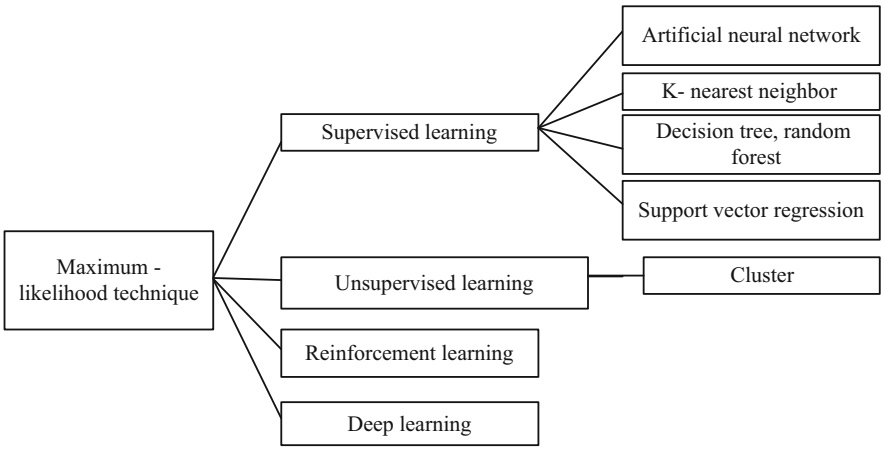


Fig. 37.11 Maximum-likelihood technique classification

may yield more accurate first-step parameter estimation by blending the advantages of each signal feature into a unified estimation framework.

In 2015, Şahin [62] studied the use of AOA–RSS hybrid information for 3D locations, owing to the limitations of the single-location method. When considering the RSS information, the AOA-based location accuracy could be improved by the least-squares method. To address the low precision of NLOS positioning, a correct filtering technology was designed and used to reduce the positioning error. In 2019, Yang [63] proposed a TOA–AOA hybrid location method based on the threshold-comparison weighting method, which used a Kalman filter to eliminate the measurement offsets generated by the signal-propagating TOA; this greatly improved the accuracy of the initial position estimation.

In 2021, Cao [64] proposed a TOA–RSS-fusion location method that used a Kalman filter, particle filter algorithm in the source range-finding phase and inertial sensors in the receiver to analyze the movement with a positioning accuracy of 2.02 cm.

The multifeature fusion location method can compensate for the non-uniqueness of a single feature and improve the location effect of the whole system; however, it increases the complexity of the system. Therefore, it is necessary to make full use of the advantages of multifeature fusion while maintaining a low system complexity.

Summary

Location accuracy, real-time performance, and robustness are three key factors of a VLP system. Current indoor location systems must improve their accuracy and anti-jamming ability based on existing theories. By modeling a complex environment, a positioning system model can be established, according to the actual environment. With the improvement in positioning accuracy, it is necessary to reduce the power

Table 37.5 Performance comparison of machine-learning-based VLP systems

Reference	Algorithm	Dimensionality	Type/size (m)	Data type	Accuracy (cm)
[57]	KNN	3D	Testing $0.35 \times 0.4 \times 0.25$	RSS	<7
[58]	WKNN	2D	Simulation $5 \times 5 \times 3$	RSS	3.4
[59]	Neural networks	2D	Simulation $5 \times 5 \times 3$	TDOA	1.662
[8]	Neural networks and WKNN	2D	Simulation $4 \times 4 \times 3$	RSS	1.5
[60]	Iterative point-by-point reinforcement learning	2D	Testing $0.12 \times 0.12 \times 0.12$	RSS	3.14
[27]	ML	2D	Testing $4 \times 4 \times 3$	RSS	3.65
[28]	Return to ML	3D	Testing $0.5 \times 0.5 \times 1.5$	RSS	1.96 and 2.16
[54]	Return to ML and ANN	2D	Testing $0.04 \times 0.04 \times 0.1$	AOA-RSS	2.6
[55]	KNN-RF	2D	Simulation $5 \times 5 \times 2.3$	RSS	1.93
[22]	ELM–KNN	3D	Simulation $4 \times 4 \times 3$	RSS	2.11
[39]	ANN, WKNN	3D	Testing $3 \times 3 \times 4$	RSS	3.18
[61]	KNN clustering and ANN	2D	Testing	RSS	90% accuracy
[29]	DT, SVR, ANN	2D	Simulation $5 \times 5 \times 3$	RSS	99.2%, 97.5%, and 96.7%, respectively
[56]	KNN, SVR, ANN	2D	Simulation $10 \times 10 \times 3$	RSS	Accuracy is greater than 95%
[26]	SVR, RF, KNN, DT	2D	Simulation $5 \times 5 \times 2.5$	RSS	SVR has maximum accuracy of 8.6

consumption and positioning system size and improve the security and privacy of users. The trends in indoor visible light positioning technology are as follows.

1. Researchers have optimized VLP systems based on the light source layout, modulation mode, and machine learning. From previous studies, it was found that a symmetrical light source has a uniform light source distribution, and the location effect is better. Modulating the light source using a multiplexing protocol

can improve the information transmission rate, and it is beneficial to distinguish the characteristic information of different LEDs. The classic geometric location method cannot satisfy the accuracy of the location system; therefore, a combination of machine learning and various intelligent optimization algorithms is the preferred choice for indoor locations.

2. The location performance of the VLP system is affected by the multipath effect. Therefore, it is necessary to reduce the noise using equalization or filtering to reduce the influence of noise on the location accuracy. It may be beneficial to design a new calibration method to improve the location accuracy of the corner and edge areas. Alternatively, a new multisensor fusion method has been studied, which is helpful for extracting multifeature signals and building a higher-precision machine-learning model.
3. Future research should focus on real-time location and tracking, 3D location, complex environment location, and microlocation systems. At present, most research is intended to localize 2D planes. Combining indoor localization systems and ML, along with the flexibility of classic and modern methods, is an important direction for realizing 3D localization. To miniaturize the system, 3D laminated-encapsulation technology must be further explored.

With the increasing demand for indoor location services, indoor location technology will be further developed. The development of new-generation communication technology will significantly promote the progress of indoor positioning technology. The application of millimeter-wave and multiple input multiple output (MIMO) technology can effectively increase the ranging accuracy and improve the anti-multipath capability of the positioning terminal. Low-cost submeter-level or even centimeter-level indoor positioning technology has become a future research trend and hot spot.

References

1. Ke Xizheng, Deng Lijun. Optical Wireless Communication [M]. Beijing: Science Press. 2016.
2. Maheepala M, Kouzani A Z, Joordens M A. Light-based indoor positioning systems: A review [J]. IEEE Sensors Journal, 2020, 20(8): 3971–3995.
3. Dawood M A, Saleh S S, El-Badawy E S A, et al. A comparative analysis of localization algorithms for visible light communication [J]. Optical and Quantum Electronics, 2021, 53: 1–25.
4. San Xinyi, Zhu Shijie, Wang Zhou, et al. Monolithically integrated light emitting and detecting micro-LED arrays and display and visible light communication applications [J]. Semiconductor Optoelectronics, 2022, 43(03): 444–450.
5. Huang Heqing, Yang Aiying, Feng Lihai, et al. Artificial neural-network-based visible light positioning algorithm with a diffuse optical channel [J]. Chinese Optics Letters, 2017, 15(5): 050601.
6. Vegni A M, Biagi M. An indoor localization algorithm in a small-cell LED-based lighting system [C]//2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, 2012: 1–7.

7. Yang S H, Kim D R, Kim H S, et al. Visible light based high accuracy indoor localization using the extinction ratio distributions of light signals [J]. Microwave and Optical Technology Letters, 2013, 55(6): 1385–1389.
8. Liu Chong, Zhang Yuexia. Research on visible light localization algorithm fusing neural network and fingerprint [J]. Semiconductor Photonics, 2019, 40(06): 891–895.
9. Kail G, Maechler P, Preyss N, et al. Robust asynchronous indoor localization using LED lighting [C]//2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014: 1866–1870.
10. Zhao Li, Han Zhongda, Zhang Feng. Research on visible indoor stereo localization based on neural network [J]. China Laser, 2021, 48(7): 0706004.
11. Do T H, Yoo M. An in-depth survey of visible light communication based positioning systems [J]. Sensors, 2016, 16(5): 678.
12. Vongkulbhisal J, Chantaramolee B, Zhao Yan, et al. A fingerprinting-based indoor localization system using intensity modulation of light emitting diodes [J]. Microwave and Optical Technology Letters, 2012, 54(5): 1218–1227.
13. Shen Rui, Zhang Jian. Indoor localization method based on visible light communication [J]. Journal of Information Engineering University, 2014, 15(1): 41–45.
14. Wu Nan, Yang Aiying, Feng Lihui, et al. Research and prospect of visible light localization key technology [J]. Journal of Nanjing University of Information Engineering (Natural Science Edition), 2017, 9(2): 159–167.
15. Rahman M S, Kim K D. Indoor location estimation using visible light communication and image sensors [J]. Int. J. Smart Home, 2013, 7(1): 99–113.
16. Kuo Yesheng S, Pannuto P, Hsiao K J, et al. Luxapose: Indoor positioning with mobile phones and visible light [C]//Proceedings of the 20th annual international conference on Mobile computing and networking. 2014: 447–458.
17. Mohammed N A, Abd Elkarim M. Exploring the effect of diffuse reflection on indoor localization systems based on RSSI-VLC [J]. Optics express, 2015, 23(16): 20297–20313.
18. Wang Xingxing, Cong Sian. A review of indoor localization research methods [J]. Software Guide, 2019, 18(09): 9–12.
19. Pang G, Kwan T, Chan C H, et al. LED Traffic Light as a communications device [C]. IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems, Tokyo, Japan, Dec., 1999, 788–793.
20. Armstrong J, Sekercioglu Y A, Neild A. Visible light positioning: a roadmap for international standardization [J]. IEEE Communications Magazine, 2013, 51(12): 68–73.
21. Sertthin C, Tsuji E, Nakagawa M, et al. A switching estimated receiver position scheme for visible light based indoor positioning system [C]//2009 4th International Symposium on Wireless Pervasive Computing. IEEE, 2009: 1–5.
22. Prince G B, Little T D C. A two phase hybrid RSS/AoA algorithm for indoor device localization using visible light [C]//2012 IEEE Global Communications Conference (GLOBECOM). IEEE, 2012: 3347–3352.
23. Cossu G, Presi M, Corsini R, et al. A visible light localization aided optical wireless system [C]//2011 IEEE GLOBECOM Workshops (GC Wkshps). IEEE, 2011: 802–807.
24. Wang T Q, Sekercioglu Y A, Neild A, et al. Position accuracy of time-of-arrival based ranging using visible light with application in indoor localization systems [J]. Journal of Lightwave Technology, 2013, 31(20): 3302–3308.
25. Guan Weipeng, Aminikashani M, Deng Peng, et al. Impact of multipath reflections on the performance of indoor visible light positioning systems [J]. Journal of Lightwave Technology, 2016, 34(10): 2578–2587.
26. Tran H Q, Ha C. Improved visible light-based indoor positioning system using machine learning classification and regression [J]. Applied Sciences, 2019, 9(6): 1–25.
27. Hsu C W, Liu Siming, Lu Feng, et al. Accurate indoor visible light positioning system utilizing machine learning technique with height tolerance [C]//Optical Fiber Communication Conference. Optica Publishing Group, 2018: M2K. 2.

28. Wu Yuchun, Chi-Wai Chow, Liu Yang, et al. [J]. IEEE Access, 2020, 8(1): 214269–214281.
29. Abdalmajeed A M M, Mahmoud M, El-Fikky A E R A, et al. Improved indoor visible light positioning system using machine learning [J]. Optical and Quantum Electronics, 2023, 55(3): 1–15.
30. Ding Deqiang, KE Xizheng, LI Jianxun. Design and simulation of light source layout for VLC system [J]. Optoelectronic Engineering, 2007, 34(1):131–146.
31. Ding Deqiang, KE Xizheng. Mathematical luminescence modeling of a general-purpose white LED [J]. Journal of Optics, 2010, 30(09): 2536–2540.
32. Zhang Xueli, Duan Jingyuan, Fu Yuegang, et al. Theoretical accuracy analysis of indoor visible light communication positioning system based on received signal strength indicator [J]. Journal of Lightwave Technology, 2014, 32(21): 4180–4186.
33. Wu Nan, Wang Xudong, Hu Qingqing, et al. A high-precision indoor visible light localization method based on multiple LEDs [J]. Journal of Electronics and Information, 2015, 37(3): 727–739.
34. Guan Weipeng, Wu Yuxiang, Wen Shangsheng, et al. Research on indoor localization technique for visible light communication based on CDMA modulation [J]. Journal of Optics, 2016, 36(11): 73–81.
35. Ye Ziwei, Ye Huiying, Nie Xiangyu, et al. A high-precision visible light localization method based on received signal strength detection [J]. China Laser, 2018, 45(3): 0306002.
36. Fang Junbin, Yang Zhen, Long Shun, et al. High-speed indoor navigation system based on visible light and mobile phone [J]. IEEE Photonics Journal, 2017, 9(2): 1–11.
37. Wang Yong, Chen Guoliang, LI Xiaoyuan, et al. A camera calibration-assisted monocular vision indoor localization method [J]. Journal of Surveying and Mapping, 2018, 9(2): 35–40.
38. Li Jingyi, Guan Weipeng. The optical barcode detection and recognition method based on visible light communication using machine learning [J]. Applied Sciences, 2018, 08(12): 1–21.
39. Chen Yirong, Guan Weipeng, Li Jingyi, et al. Indoor real-time 3-D visible light positioning system using fingerprinting and extreme learning machine [J]. IEEE Access, 2019, 8: 13875–13886.
40. Zhang Rui. Research on indoor localization algorithm based on visible light communication [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2021.
41. Qin Ling, Zhang Chongtai, Guo Ying. Research on visible light indoor localization algorithm based on Elman neural network [J]. Journal of Optics, 2022, 42(5): 0506002.
42. Xu Wei, Wang Jia, Shen Hong, et al. Indoor positioning for multiphotodiode device using visible-light communications [J]. IEEE Photonics Journal, 2015, 8(1): 1–11.
43. Shen Rui, Zhang Jian, Wang Ding. Indoor localization algorithm based on visible light communication and the corresponding parameter estimation of the Claramello boundary [J]. Advances in Lasers and Optoelectronics, 2014, 51(9): 85–92.
44. Zhao Li, Peng Kai. Optimization of indoor visible light communication light source layout based on white light emitting diode [J]. Journal of Optics, 2017, 37(7): 49–57.
45. Guan, Shuaitang. Characteristics of Light Source Layout and Realization Mechanism of Indoor Visible Light Positioning System [D]. Beijing: Beijing University of Posts and Telecommunications, 2021.
46. Luo Junhai, Fan Liying, Li Husheng. Indoor positioning systems based on visible light communication: State of the art [J]. IEEE Communications Surveys & Tutorials, 2017, 19(4): 2871–2893.
47. Kang Ye, Ke Xizheng. Multidimensional coding in visible optical communication [J]. China Laser, 2015, 42(02): 138–144.
48. Xizheng Ke, Zhiyun Yin. Coding theory in wireless laser communication systems [M]. Beijing: Science Press, 2009.
49. Ke, Xi-Zheng. Principles and Applications of Wireless Optical Orthogonal Frequency Division Multiplexing [M]. Beijing: Science Press. 2017.
50. Li Mengfan. Research on Optical Wireless Communication system with multi-band CAP modulation [D]. Xi'an: Xi'an University of Technology, 2017.

51. Bawazir S S, Sofotasios P C, Muhaidat S, et al. Multiple access for visible light communications: Research challenges and future trends [J]. Ieee Access, 2018, 6(2): 26167–26174.
52. Xiang Yang. Research on key technology of MIMO-OFDM for visible optical communication [D]. Beijing: Beijing University of Posts and Telecommunications, 2015.
53. Tran H Q, Ha C. Machine learning in indoor visible light positioning systems: A review [J]. Neurocomputing, 2022, 49(1): 117–138.
54. Hong Chongyou, Wu Yuchun, Yang Liu, et al. Using silicon photovoltaic cells and machine learning and neural network algorithms for visible-light positioning systems [J]. Optical Engineering, 2020, 59(9): 096107–096107.
55. Tran H Q, Ha C. Fingerprint-based indoor positioning system using visible light communication—a novel method for multipath reflections [J]. Electronics, 2019, 8(1): 1–16.
56. Dang Shuping, Ma Guoqing, Shihada B, et al. Enabling smart buildings by indoor visible light communications and machine learning [J]. Electronics, 2019, 2(5): 1–7.
57. Xu Ming, Xia Weiwei, Jia Ziyan, et al. A VLC-based 3-D indoor positioning system using fingerprinting and K-nearest neighbor [C]//2017 IEEE 85th Vehicular Technology Conference (VTC Spring). IEEE, 2017: 1–5.
58. Van M T, Van Tuan N, Son T T, et al. Weighted k-nearest neighbour model for indoor VLC positioning [J]. Iet Communications, 2017, 11(6): 864–871.
59. Li Xiaoji, Cao Yaru, Chen Chen. Machine learning based high accuracy indoor visible light location algorithm [C]//2018 IEEE International Conference on Smart Internet of Things (SmartIoT). IEEE, 2018: 198–203.
60. Zhang Zhuo, Zhu Yaquang, Zhu Wentao, et al. Iterative point-wise reinforcement learning for highly accurate indoor visible light positioning [J]. Optics Express, 2019, 27(16): 22161–22172.
61. Munir B, Dyo V. Passive localization through light flicker fingerprinting [J]. IEEE Sensors Journal, 2019, 19(24): 12137–12144.
62. Şahin A, Eroğlu Y S, Güvenç I, et al. Hybrid 3-D localization for visible light communication systems [J]. Journal of Lightwave Technology, 2015, 33(22): 4589–4599.
63. Yang Chaochao, Chen Jianhui, Liu Deliang, et al. Joint indoor localization algorithm based on arrival time and arrival angle [J]. Advances in Lasers and Optoelectronics, 2019, 56(22): 222802.
64. Cao Yang, Dang Yuchao, Peng Xiaofeng, et al. TOA/RSS hybrid information indoor visible light localization method [J]. China Laser, 2021,48(1):133–141.

Research Progress on Visible-Light Communication Uplinks

38

Contents

Introduction 1436
Visible-Light Communication Systems and Uplinks 1437
Visible-Light Communication-System Composition 1437
Visible-Light Communication Links 1437
Visible-Light Communication Uplinks 1438
Uplink Schemes for Visible-Light Communication 1439
Infrared Communication Uplinks 1440
Domestic and Foreign Research Progress on Uplinks 1442
RF Uplinks 1448
Composition of an RF–VLC Two-Way Communication System 1448
Uplink RF Communication-Link Channel 1450
WiFi–VLC Communication System 1450
Power-Carrier Uplinks 1452
PLC–VLC-Fusion Communication-System Structure 1452
Channel Model of a PLC–VLC-Fusion Communication System 1453
PLC–VLC-Fusion Communication-System Scheme 1453
Uplinks with Reverse-Modulation Technology 1455
Structure of a Reverse-Modulation System 1456
Reverse-Modulation Uplinks 1456
Summary 1458
References 1459

Abstract

Uplink refers to the link through which the response signal is transmitted from the receiving end to the sending end. This paper summarizes the research progress of visible light communication uplink at home and abroad, and the existing application scheme of LED visible light uplink, mainly describes the principle and scheme of RF uplink, visible light uplink, infrared uplink, power carrier uplink, and visible light uplink using reverse modulation technology. Finally, the advantages and disadvantages of existing schemes are given, and the development direction of visible light communication uplink is forecasted.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_38

Keywords

Uplink · Visible light communication · Reverse modulation · Infrared light communication

Introduction

Visible-light communication (VLC) is a type of optical wireless communication. If light-emitting diodes (LEDs) are used as the baseband signal, the light wave emitted by the white LED can be modulated so that it not only transmits light, but also transmits data [1]. As LEDs are gradually becoming the main lighting devices in the world, VLC technology has also attracted the attention of researchers, resulting in its rapid development.

Visible-light communication can solve the shortcomings [2] of traditional RF radio communication. That is, it has the advantages of a simple system structure, large communication capacity, no required spectrum licensing, and no electromagnetic radiation. It is harmless to the human body and is not subject to electromagnetic interference. It has high privacy and confidentiality, low cost, and green environmental protection. These advantages enable visible-light communication technology to be well applied in various scenarios, as shown in Table 38.1.

An uplink refers to the transmission path from a terminal device to the network. Uplinks are mainly proposed to solve the problem of network congestion and are used extensively in today's networks. Downlinks are prone to congestion, resulting in a slow network speed, while uplinks are relatively stable. In addition, a complete communication system must contain a complete two-way communication link, that is, it must include uplinks and downlinks. Downlink application technology has matured, but uplinks have consistently been a problem in two-way communication systems. Existing visible uplink schemes are mainly for radio frequencies, infrared light, and visible light.

This chapter provides a brief overview of the research status of LED visible-light communication uplinks at home and abroad, and introduces five types of uplinks: RF, visible light, infrared, power-carrier using reverse modulation, and indoor

Table 38.1 VLC communication application scenarios

Communication range	Application scenarios
Indoor environment	Supermarket and factory locations Smart home, Internet of Things (IoT) Encrypted communications Hospitals, gas stations, ammunition depots, and other places with electromagnetic limitations
Outdoor environment	Vehicle-to-vehicle (V2V) and Internet-to-vehicle (I2V) intelligent transportation Smart city-based lighting facilities

visible-light. Existing schemes, principles, advantages, and disadvantages are also presented.

Visible-Light Communication Systems and Uplinks

Visible-Light Communication-System Composition

The basic structure of a visible-light communication system is shown in Fig. 38.1. It is mainly composed of sending, transmitting, and signal-receiving modules. The system transmission information flow is as follows.

First, to meet the communication requirements, the binary sequence to be sent is encoded and modulated. The modulated electrical signal drives the LED, and the LED modulates the message to be sent into an optical signal. The photoelectric detector receives the optical signal transmitted on the channel, and then transforms the optical signal into an electrical signal. The electrical signal is weak and contains high-frequency noise; therefore, signal processing is necessary to increase the signal and a low-pass filter is used to filter the noise to obtain the appropriate electrical signal. Finally, the signal is demodulated and decoded to restore the binary sequence [3].

Visible-Light Communication Links

Visible-light communication channels differ in different application scenarios. In an outdoor visible-light communication system, the system generally adopts a free-space optical (FSO) channel model [4], as shown in Fig. 38.2; when the outdoor visible light communication is carried out, the noise introduced by the complex outdoor environment and complex background light in the free space will affect the VLC system and cause the light to attenuate.

The line of sight (LOS) indicates the range within which the transmitter and receiver in a visible-light communication system can "see" each other; that is, the transmitter and receiver are within each other's field of view (FOV). For example, in an indoor visible-light communication system, the system usually adopts a direct or diffused line-of-sight communication link [1].

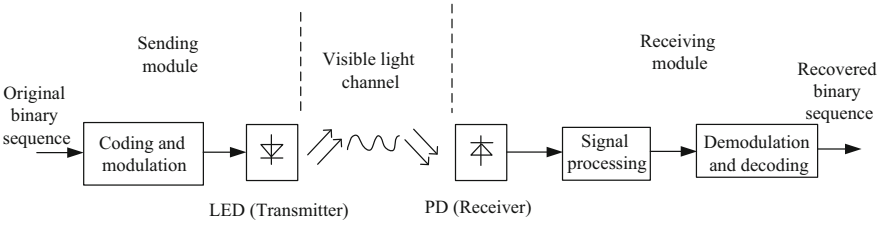


Fig. 38.1 Structure of a visible-light communication system



Fig. 38.2 Schematic diagram of an FSO system

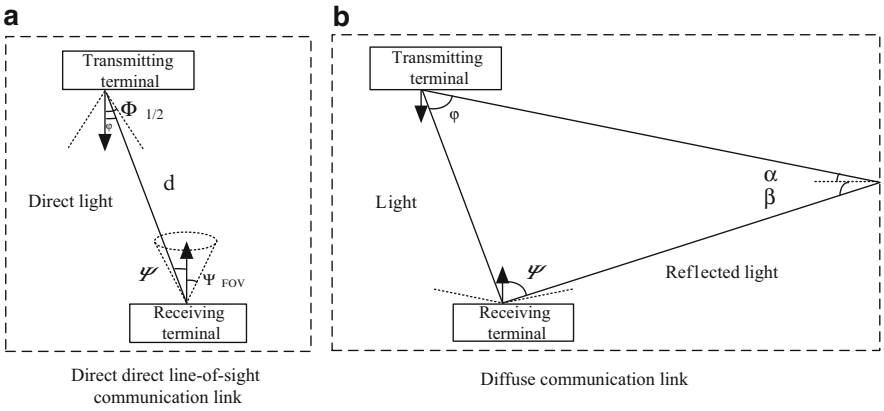


Fig. 38.3 Two types of indoor visible-light communication links

In a direct LOS link, the FOV angle between the transmitting and receiving ends is small. The two ends must be aligned with each other when communicating, and no obstacles can be allowed to block the communication link, as shown in Fig. 38.3a.

In a diffuse communication link, the transmitting and receiving ends of the FOV angle are usually relatively large, with the light evenly distributed in the room. The receiving end can receive direct light from the transmitting end and reflected light from a wall, as shown in Fig. 38.3b. In a diffuse communication link, although the FOV angle is large, the transmission process will produce a multi-path effect [1].

Visible-Light Communication Uplinks

The schematic diagram of a visible-light communication uplink is shown in Fig. 38.4. In a room, the LED hanging from the ceiling serves as the transmitter end of the communication system, and the user's personal computer (PC) serves as the receiver end. The uplink refers to the transmission path from the terminal device (mobile phone, computer, etc.) to the network, while the downlink is the transmission path from the network to the terminal device.

In the downlink, the sequence of information to be sent is first coded and modulated to produce an AC electrical signal. The optical signal emitted by the

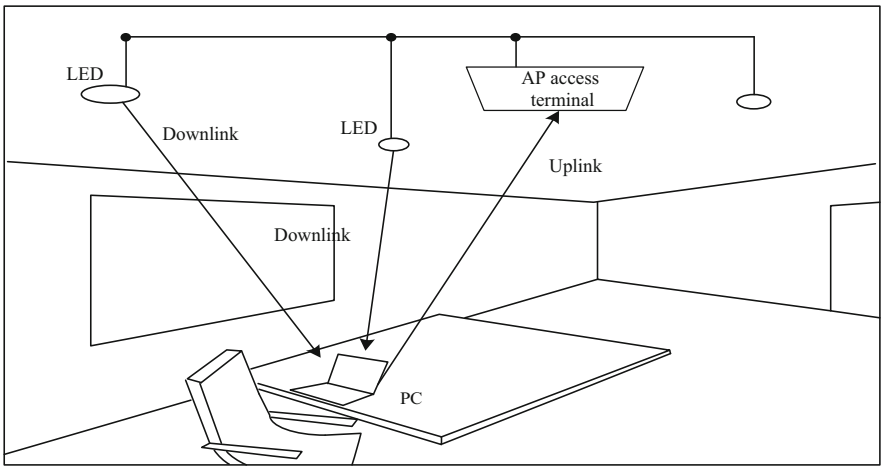


Fig. 38.4 Schematic diagram of an uplink and downlink

LED is detected by the photodetector at the receiving end through the downlink, and the optical signal is converted to an electrical signal. Finally, the original information sequence is recovered from the electrical signal.

In the uplink part, the response signal is sent by the LED driver using visible-light communication, and the signal is received by the photoelectric detector at the LED sending end, forming a closed uplink-communication link.

Uplink Schemes for Visible-Light Communication

Scheme 1: Reference [5] proposes and implements a system that uses an LED as the transceiver component of the optical communication system and can conduct real-time bidirectional visible-light communication. Both the upstream and downstream links of the scheme are visible light, and the system adopts hardware to rapidly switch between the system modes. The system adopts a time-division multiplexing (TDM) method to enable the LED to work in half-duplex mode and designed a communication-control protocol to realize the bidirectional control of the VLC system. An experimental device platform was constructed to realize bidirectional visible-light communication at a distance of more than 2 m and a rate of 8 Mbit/s.

Figure 38.5 shows how to achieve visible-light communication using a TDM channel. When two LED devices send data to each other, they occupy the channel respectively through the upstream and downstream links of different time slots for data transmission. Only one LED can transmit in the same time slot; the other LED is in the sensitive state to receive the sent data. Through channel segmentation, the LEDs can conduct half-duplex bidirectional communication.

Scheme 2: Reference [6] proposed to use an RGB-LED (red, green, and blue light-emitting diode) as the light-emitting element of the system. The blue light was loaded

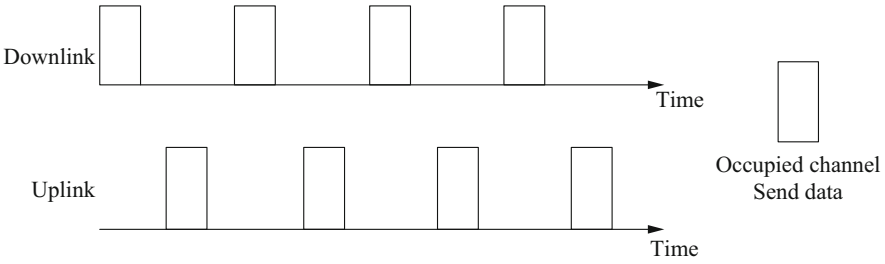


Fig. 38.5 Principle of time-division multiplexing communication in uplinks and downlinks

with a 32/64 orthogonal-amplitude modulation–orthogonal frequency-division multiplexing (32/64QAM-OFDM) uplink signal as the uplink channel of the system.

Loading 32/64QAM-OFDM downlink signals onto the red and green light channels achieved a transmission rate of 300 Mbit/s on the uplink in the 66 cm free space. By installing a filter on the front of the detector, the system can effectively suppress the mutual interference between different wavelength channels and form a wavelength-division duplex system in the visible band. The three-color LED used by the wavelength-division multiplexing technology has only three fixed channels; hence, it cannot flexibly allocate the uplink and downlink rates in the time-division multiplexing method.

Infrared Communication Uplinks

An uplink for infrared light is similar to that for visible light. As shown in Fig. 38.6, general infrared light adopts 780–950 nm wavelengths to transmit information. In early infrared communication technology, the main band of 780–950 nm light had many advantages, such as a small path loss, long-distance transmission, and a narrow infrared-light spectrum with the same emission power.

A narrowband infrared system has a better ability [7] to resist environmental noise interference than a visible light communication system. The 780–950 nm band is exactly consistent with the silicon detector peak frequency band; therefore, the advantages of high infrared detection efficiency in this band make infrared light suitable for use in uplink implementation schemes. Table 38.2 shows a comparative analysis of infrared and white LED communication in a room.

However, infrared communication also has disadvantages. Human eyes are sensitive to 780–950 nm infrared light; hence, the transmitting power must be strictly limited. Because an infrared light beam is more concentrated than visible light, it is necessary to aim more accurately when building the system. Therefore, the link quality of infrared communication will be more limited than that of visible-light communication.

In [6], Cossu et al. used an infrared LED at the sending end and an APD (avalanche photo diode) detector at the receiving end. They placed 473 nm and

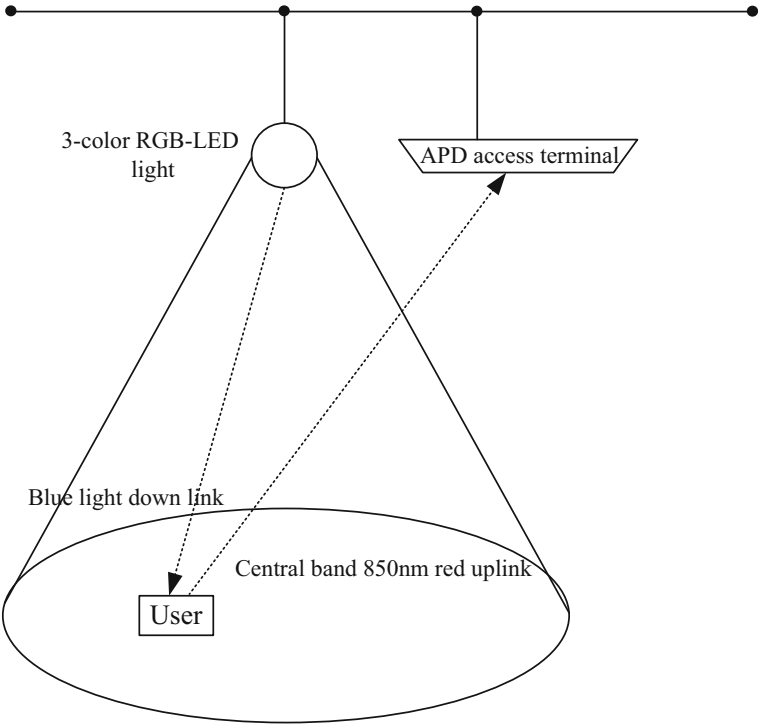


Fig. 38.6 Schematic diagram of an infrared uplink

Table 38.2 Comparison and analysis of infrared and white light in indoor wireless communication

Typical characteristics	Infrared light	White light
Signal light source	Infrared LED	White LED
Working wavelength	800–900 nm	380–780 nm
Modulation bandwidth	Tens of kHz–hundreds of MHz Tens of kHz–hundreds of GHz	Tens of kHz–hundreds of MHz
Interior layout	Additional infrared-communication light source and route must be built	Indoor circuit layout is simplified.
Shadow effect	Susceptibility to shading	Multiple LEDs can be placed to eliminate shadow effects
Transmit power	Transmit power must be limited	No limits in general

805 nm long-wave pass filters before the receiving end and the sending end, respectively, to realize a wavelength-division duplex system. The front filter effectively removed the interference between the upper and lower lines and suppressed the background noise.

The system uplink used an infrared light with an 850 nm central wavelength for communication, the downlink used the blue light channel in an RGB three-color LED, and the upper and lower links used discrete multi-tone (DMT). At the end of the experiment, 400 Mbit/s two-way optical wireless communication at a 2 m distance under 500 lx ambient light was realized.

Domestic and Foreign Research Progress on Uplinks

To address the difficulty of realizing an all-optical communication link in a visible-light communication system, Ke et al. [8] designed a scheme based on defocusing cat's-eye reverse-modulation technology. The upper and lower links of this scheme were independent of each other. They adopted quaternary frequency-shift keying and differential quadratic phase-shift keying modulation to solve the problem that the transceiver ends of a traditional single-light-source full-service communication system needed to cooperate with each other. Finally, in a 3 m data-transmission link, the downlink-transmission rate was 5 Mbit/s, and the uplink transmission rate was 2 Kbit/s. This was the first reported uplink in the field.

Qiu et al. [9] proposed a full-duplex reverse-modulation system using a single light source. The full-duplex system had two modulation processes. To achieve better full-duplex communication, the system proposed that the modulated on–off keying (OOK) signal generated by the signal generator to the light source of the optical transmitter be larger than the amplitude-modulation (AM) signal generated by the signal generator to the Modulated Retro-Reflector (MRR). This view was verified by simulation. Finally, an indoor close-range single-light-source full-duplex reverse-modulation experiment was conducted to achieve a modulation rate of 2.048 Mbit/s at the inquiry end and 2.0 Kbit/s at the reverse end.

Ding et al. [10] designed a portable optical antenna device for full duplex communication. The device uses a single chip microcomputer as the master controller and a 1W white LED and photodiode (PIN) for transmission and reception in the upper and lower communication links, respectively. The system adopted Manchester encoding for encoding and decoding and a serial universal asynchronous receiver/transmitter (UART) interface. They finally realized an up-and-down transmission rate of 115.2 Kbit/s over a 3 m transmission distance for real-time full-duplex communication data reception.

Liu et al. [11] experimentally demonstrated bidirectional transmission links using LED visible-light communication for both downlink and uplink paths. They proposed and demonstrated that time-division duplexing (TDD) could significantly eliminate reflection interference in visible-light communication. A simple OOK control was used to realize 2 m free-space bidirectional transmission.

Using a single light source, Xu et al. [12] successfully built a duplex communication system by combining internal and external modulation. The system uses internal modulation for uplink communication and acousto-optical external modulation for downlink communication, in which the downlink–uplink communication rate is up to 30 Kbit/s, and the downlink communication rate can reach up to 3 Mbit/s.

Little et al. [13], from Boston University, proposed an indoor-fusion lighting system that combined WiFi and VLC. The system used a downlink VLC channel to supplement the uplink RF communication and used a switching mechanism between the uplink WiFi and downlink VLC to dynamically allocate resources and optimize the system throughput. Experiments showed that the communication rate of the VLC downlink reached 10 Mbit/s. The communication rate of the WiFi uplink reached 30 Mbit/s.

Basnayaka [14] of the University of Edinburgh, UK, designed a hybrid system consisting of multiple VLC and RF access points. To improve the interrupt rate and coverage of each user of an independent VLC network, a simple RF-deployment method was proposed. The paper also quantified the minimum spectrum and power requirements of the RF system. As demonstrated through simulation, the method greatly improved the coverage of an indoor visible-light communication system.

Lu et al. [15] designed a full-duplex visible-light communication system that met the technical requirements of Ethernet transmissions. The system was composed of a near end and a far end. The near end used Ethernet optical cat as a terminal for sending and receiving, and the far end was connected to household equipment for transmission. The system downlink was transmitted from the near end to the remote home device and the uplink was sent from the remote home device to the near end of the interface, to complete the full-duplex communication. The upstream and downstream link signals went through the differential to the single-ended circuit, from the single-ended to the differential circuit, then from the single-ended to the differential circuit into the interface into the remote home device. Experiments showed that the system could realize 100 Mbit/s full-duplex Ethernet communication when the communication distance was 10 m.

The Disney Research Institute and Schmid et al. [16] at ETH Zurich proposed a novel approach for using VLC as a short-range, low-bit-rate wireless ad-hoc network for uplinks, downlinks, and visible-light communication (VLC). The system achieved data rates of 800 Mbit/s at a distance of 2 m. New designs for a low-complexity VLC physical layer and VLC media-access-control protocol layer were introduced, and the effectiveness of these two layers was evaluated.

Solemn et al. [17] discussed in detail the application of the visible light communication uplink scheme and proposed that radio frequency, visible light, 780–950 nm red light, and laser could be used in a visible-light uplink scheme. They presented the advantages and disadvantages of each scheme and gave the prospects. They also designed a new VLC uplink scheme using a 1550 nm laser as the carrier. The scheme used a wavelength-division duplex uplink for the 1550 nm laser visible-light downlink. Simulation results showed that, after a preliminary analysis, the uplink rate could reach 1 Gbit/s, which could be matched by the visible-light downlink rate.

Gfeller [18] established an indoor infrared communication-system model, divided indoor channels into line of sight (LOS) and non-line of sight (NLOS), and calculated their impulse response to reflect the characteristics of indoor channels. The indoor infrared communication system was simulated and analyzed.

Zhang et al. [19] proposed a full-duplex radio-over-free-space-optical (RoFSO) transmission-system architecture that physically preserved the uplink light source. The optical center carrier of the downlink double-sideband (DSB) modulated signal was extracted as the uplink light source, which reduced the system complexity.

Komine et al. [20] studied the influence of shadows and wall multipath reflections on the system in a visible-light channel transmission by adopting a time-division multiple access (TDMA) modulation mode. They obtained a data-transmission rate of 800 Mbit/s on the uplink and effectively reduced the influence of shadows on the system.

Yamazato et al. [21] introduced a feasibility study of an uplink VLC beacon system for a universal traffic-management system (UTMS) and proposed an uplink V2I system for UTMS. The system used a photodiode (PD) as a VLC receiver and a commercially available LED headlight as a transmitter. Multiple PDs with lenses were found to be effective, both for forming small communication areas and for applying bit-loading algorithms to further increase the data rate. The results achieved a throughput of 3.1 Mbit/s.

Rana et al. [22] targeted unmanned aerial vehicle–Internet of Things (UAV-IoT)-based communication networks. They proposed an uplink non-orthogonal multiple access (NOMA) resource-allocation algorithm, which provided services for IoT devices of different applications. The optimization problem was formulated under constraints, with the goal of maximizing the data rate and minimizing latency. The scheduler worked in the time domain to optimize the delay and in the frequency domain to optimize the data rate. Power allocation was used to ensure a fair data rate between users who were assigned the same resource block. Simulation results showed that the algorithm significantly improved the fairness, latency, and reliability of the system, while achieving a spectral efficiency and summation rate close to the upper limit of the system.

Alicja et al. [23] proposed a software-defined narrowband Internet of Things (NB-IoT) uplink framework in the design field, as well as its implementation and potential use cases. The framework, which could be used as an academic tool for developing, researching, and optimizing digital-transmitter paths, also highlighted the need for data-processing optimization to minimize the power consumption and resource usage of NB-IoT nodes when transmitting collected telemetry data.

Nakamura et al. [24] proposed cooperative serial-interference cancellation (C-SIC), which solved the problem of adopting uplink NOMA in multi-cell networks. When the uplink could reach multiple base stations, the signal must complete the link SIC at each base station, which reduced the network throughput. The final conclusion was that the SIC processes of different base stations were transmitted through a backhaul link through a combination of signals. Compared with conventional NOMA, it was found that C-SIC could promote the application of uplink NOMA to multi-cell networks.

Kim et al. [25] proposed a spatial-reuse method for uplinks, which used basic-service-set (BSS) color and proximity information to improve the efficiency of carrier perception and thus improved the spatial reuse. Simulation results showed that proximity-based sensitivity control for uplinks (PSC-UL) was comparable to

overlapping BSS/preamble detection on traditional OBSS/PD. The overlapping basic service set/preamble-detection uplink method considerably improved the throughput.

Rosenthal et al. [26] proposed a differential quadrature phase-shift keying (DQPSK) backscatter uplink for a neural brain–computer interface (BCI) as an alternative to active radio. The uplink achieved a throughput of 25 Mbit/s when operating in the 915 MHz industrial, scientific, and medical (ISM) band.

Cahyadi et al. [27] experimentally demonstrated an indoor uplink near-infrared LED camera communication mechanism called an Internet call center (ICC). It used near-infrared (NIR) light as the communication medium and a camera as the receiver, which differed from past studies, which only focused on positioning research. The ICC provided a practical positioning method for uplink indoor wireless communication. It used camera receivers to provide wider coverage and accurate indoor positioning in IR communications.

Kandeepan et al. [28] proposed an exponential waveform modulated optical wireless communication (OWC) uplink where each waveform was treated as a virtual transmitter. Additional data were transmitted via the index of the selectively activated virtual transmitter, and increased the OWC uplink speed while achieving similar bit-error rate (BER) performance. The exponential waveform modulated uplinks were experimentally demonstrated in NIR OWC systems. The results showed that the proposed waveform-index modulated uplink could achieve a data rate of up to 15 Gbit/s, which significantly improved the data rate under similar power and BER conditions.

Yun et al. [29] analyzed a VLC link, the characteristics of the VLC channel, and visible-light demodulation technology in detail. They simulated an OOK modulation and demodulation scheme and an orthogonal frequency-division multiplexing (OFDM) modulation and demodulation scheme. The simulation results showed that both technologies could be used as indoor visible-light modulation technologies. To reduce the influence of intersymbol interference (ISI) caused by high-speed data transmissions, the OFDM modulation and demodulation scheme was selected, which could be used for indoor lighting and high-speed data transmissions. Finally, an implementation scheme for VLC modulation and demodulation was given, which laid a theoretical foundation for further experimental research.

Ke et al. [30] from Xi'an University of Technology proposed a multi-dimensional coding method of a visible light communication system with a multi-primary LED modulated by PPM, which could effectively improve the spectrum utilization rate and small time-slot error rate of visible light communication. Finally, the hardware platform of an indoor visible light communication system with an RGB3-color LED was constructed, and a multi-primary LED was used for multi-dimensional coding. The experiment realized 1.5 Mbit/s video communication at a 3 m distance.

Song et al. [31] proposed a new cost-effective single-frequency network (SFN) video-broadcast system-design scheme based on the deep integration of a programmable logic controller (PLC) and VLC. This scheme greatly reduced the complexity of the VLC network protocol and provided better information coverage. In the

following year, they studied the PLC-VLC system channel problem, and analyzed the PLC-VLC system channel as a whole for the first time.

Chen et al. [32] detailed the current research status, key technologies, application scenarios and development trend of visible light communication with white LEDs, in which the literature gives that the future visible light communication must contain a complete two-way communication link, and gives two scenarios in which the RF uplink and VLC form a converged system with power carrier and RF.

Ke et al. [33] comprehensively introduced visible-light LED communication technology, the basic principle of LED-device luminescence, the optimal-position deployment of LEDs in an indoor VLC system, the array design at the receiving end, and three link-configuration methods at the receiving and sending ends of the VLC system. Moreover, they presented the bottleneck faced by VLC-technology research institutes.

Zhang et al. [34] proposed a full-duplex communication system using a dual-wavelength emission mode. The system adopted a 1530-nm laser-communication light source in the forward link and 1550 nm laser-communication light source in the reverse link. The optical signals between the forward and reverse links did not interfere with each other, which effectively improved the communication rate of the reverse link. Moreover, the system simplified the complexity of the full-duplex MMR communication system. The simulation results showed that the communication rate reached 1 GHz and the communication distance was 300 km. The link communication margin was greater than 5 dB, and the bit-error rate was better than 10^{-15}.

Duan et al. [35] proposed an uplink signal transmission in the general model of a multiple-input multiple-output (MIMO) system to remove cellular mass. Based on an analysis of the TDD and frequency-division duplex (FDD) modes, the source of pilot pollution in the system running in TDD mode and the classification of the effect function in power optimization were observed. Using partial minimum mean-square error (MMSE) signal-detection technology and access-point selection technology, the spectrum-efficiency expression of the system uplink was derived. To improve the system's spectral efficiency, the pilot-allocation algorithm was designed first, and then the pilot-power and data-power maximum–minimum optimization problem was modeled.

Liu et al. [36] studied a type of Ethernet mutual-integration technology for LED VLC. The technology integrated LEDs and Ethernet through an optical-communication medium-conversion module. The system used wired media as the downlink and wireless media as the uplink. Finally, an experiment on mutual-convergence network communication was conducted. The experimental results showed that the data transmission between each client in the fusion local area network (LAN) could be very good, and the communication rate could reach 1 Mbit/s.

Sun et al. [37] proposed and designed an Ethernet communication system to solve the problem of transmitting uplink and downlink Ethernet signals without

modulation. They built a test environment to conduct a 10 m distance network-communication test with both uplink and downlink transmission rates of 100 Mbit/s.

Based on the original indoor VLC network, Feng et al. [38] used a combination of VLC and laser communication to realize an indoor wireless-communication LAN. They considered whether the uplink should be a laser instead of visible light, or whether visible light should be used for the downlink. The system optical code-division multiple access (OCDMA) technology used optical orthogonal code (OCC) as address codes to encode and transmit signals. The performance of the system before and after the improvement was compared and analyzed. The simulation results showed that the system performance was obviously improved, compared with the traditional system where both the upper and lower links are visible.

Zeng et al. [39] discussed the application of VLC and related business scenarios of telecom operators. They proposed a fusion system using visible-light and base-station back-transmission; that is, radio-frequency propagation was used in the downlink, and visible light was used in the uplink. The transmission performance of the fusion system was analyzed. In the R8 version of a long-term evolution (LTE) network, the system bandwidth was 20 MHz. The downlink rate was 100 Mbit/s and the uplink rate was 135 Mbit/s.

Zhu et al. [40] proposed the use of relay technology in VLC, which could solve the problem of link occlusion in VLC technology. The role of a relay in the system differed in the upper and lower links. The relay node in the downlink was mainly to assist the source node to forward information. The role of the relay node in the uplink was to reduce the transmission power of the light source and reduce the damage to the human eye.

Wang et al. [41] proposed a new uplink-scheduling algorithm that solved the problems of poor fairness and high packet-discarding. The algorithm established a target integer linear-programming model according to the delay constraints of real-time services, and then scheduled according to the target integer linear-programming model. Simulation experiments were conducted in which the method was compared with the dynamic Hungarian algorithm with modification (DHAM). The performance fairness of DHAM was improved without sacrificing the MAC-layer throughput.

To improve the uplink, Wu et al. [42] studied the user-association and inter-cell interference problems. For the user-association problem, the advantages and disadvantages of each method were presented, and a new uplink-interference mitigation technology was proposed, which mitigated the co-channel interference through a dynamic distributed power-control scheme. To solve the inter-cell interference problem, an improved inverse frequency-multiplexing scheme was proposed. The purpose of this scheme was to solve the spectrum-allocation problem, so that the uplinks of users in densely distributed areas could be distributed to different frequency bands for uplink communication. This simultaneously increased the distance of users in the same frequency, reduced the interference of users in the same frequency, and increased the utilization rate of the spectrum. Finally, the effectiveness of the dynamic distributed power-control scheme and the improved inverse frequency-multiplexing scheme was verified by simulation.

Komine et al. [43, 44] proposed a fusion system, then presented the overall structure of the fusion system, and conducted a qualitative analysis of power-line communication (PLC) and VLC channels.

Amirshahi et al. [45, 46] evaluated the performance of integrated power-line and visible-light communication systems using turbo-coded OFDM. They analyzed PLC channels, VLC channels, and channel noise.

In 2015, Song et al. [47, 48] analyzed the integrated channel of a fusion system, and adopted a series PLC channel, VLC channel, LED equivalent channel, and other forms to create a fusion-system channel-analysis model [49].

RF Uplinks

Wireless-network data consumption is growing dramatically, owing to the growing demand for mobile services and applications. RF communications using deployed networks are characterized by shared media, limited available spectrum, and limited ability to scale. Directional communications, including visible-light communications, can limit contention in controlled environments. If properly distributed, they can provide massive aggregated data capacity [50–52] in indoor scenarios. As Internet-Protocol (IP) data traffic tends to increase wireless usage and video streaming, the limitations of deployed RF technology will become apparent.

Traditional wireless-communication systems use radio frequencies as the carrier for information transmission; however, visible-light communication can still use radio frequency as the uplink [53–55]. To solve the VLC uplink problem in the physical layer (PHY), researchers have studied other heterogeneous schemes, such as power-line communication (PLC) and visible-light communication (VLC) fusion systems, WiFi and visible-light communication fusion systems [56], etc.

Visible-light communication has advantages that radio-frequency communication does not; however, it is not without disadvantages. In outdoor environments, owing to environmental and sun lighting, visible-light communication systems will be subjected to background noise. Moreover, with indoor VLC, because light cannot penetrate walls, each room in a building will need to set up access points to ensure that it can use visible-light communication. Table 38.3 [57] shows a comparison of the main features of radio-frequency and visible-light wireless communication systems.

In summary, both VLC and RF communication have advantages and can be used as alternative uplinks. An optical wireless communication system is more suitable for indoor applications with short distances, high throughput, and low mobility, while RF-wireless communication systems are suitable for long-distance environments with high mobility.

Composition of an RF–VLC Two-Way Communication System

The sending and receiving ends of a hybrid RF–VLC visible-light bidirectional communication system are shown in Fig. 38.7. The complete system includes a downlink visible-light communication link and uplink RF communication link.

Table 38.3 Comparison of the main characteristics of optical and radio-frequency wireless communication systems

Media characteristics	Radio frequency	Optical communication
Bandwidth is limited	Yes	No
Can traverse through walls	Yes	No
Multipath fading	Yes	No
Multipath distortion	Yes	Yes
Path loss	High	High
In-band noise source	Other-user interference	Background noise

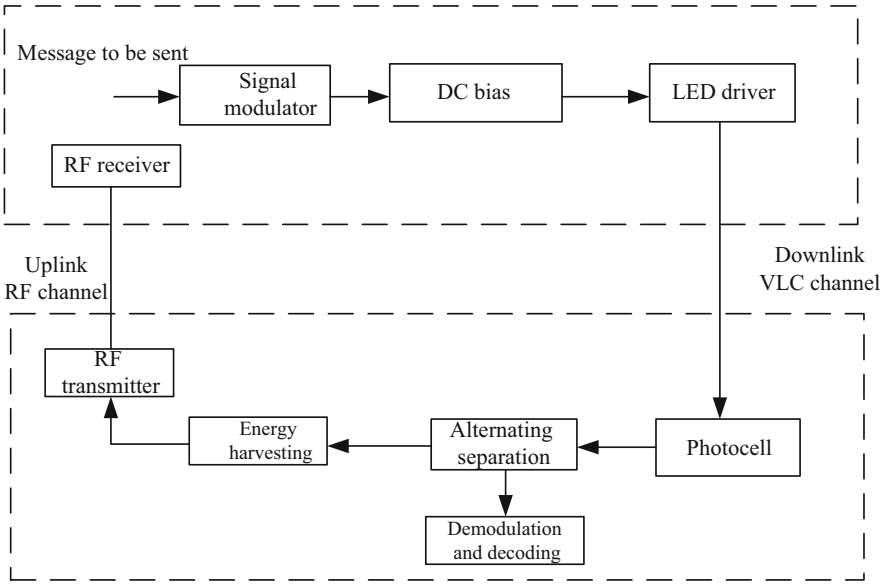


Fig. 38.7 Block diagram of an RF–VLC system

On the downlink, the transmission information sequence is coded and modulated to generate AC electrical signals, and then a DC bias is added to ensure that the output signal of the drive LED is non-negative. The LED emits a light signal through a direct line-of-sight channel using the receiving end of the photocell detection; then, the light signal is converted into an electrical signal that is separated into DC and AC components by AC and DC separation. The original information sequence is restored by the AC component and the DC component is used for energy collection.

The uplink part uses the electrical energy collected during visible-light communication to drive the RF transmitter to transmit the response signal and is received by the RF receiver at the LED sending end to form a closed uplink communication link.

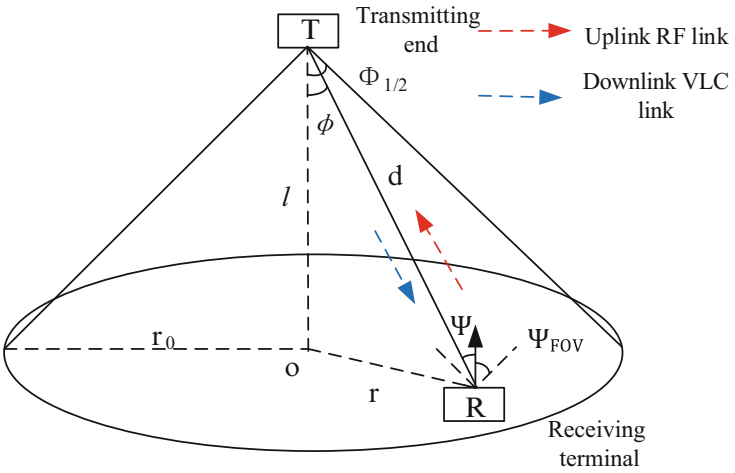


Fig. 38.8 Downlink VLC and uplink RF communication links

Uplink RF Communication-Link Channel

Figure 38.8 shows how the system achieves bidirectional communication in an indoor environment. We assume that there is only one pair of sending and receiving ends in the entire system. The sending end T is in the ceiling, the power half-angle of the LED light source is one-half, and the LED can cover the communication range delimited by the cone in the figure, with height l and bottom radius r_0.

The receiving end R can occur randomly at any place on the bottom of the cone, and it is distance r from the center O of the bottom surface. In the communication process, the sending end first sends the message to the receiving end through the downlink. When the receiving end successfully receives the signal, the collected electrical energy is used to drive the RF device to send the response signal, which forms an overall two-way communication system.

WiFi–VLC Communication System

At present, WiFi uplinks have become the mainstream of RF uplinks. The omnipresent and omnidirectional characteristics of WiFi can be easily used as uplinks; hence, WiFi and VLC fusion communication systems have become mainstream, especially if VLC can be used to reduce the congestion on an RF uplink as a heterogeneous network. Indoor wireless local-area networks (WLANs) are currently the most widely used and deployed, and the most mature technology is the WiFi protocol, that is, the IEEE 802.11-series protocol.

Existing mobile terminals have more installed WiFi modules, and WiFi hotspots have been generally deployed; however, with the increasing number of users, the collision probability and the waiting time of retreating into the

channel are also increasing. Consequently, the visible-light communication system can be used to carry downlink data, and WiFi is used for uplink data. Visible-light communication is a line-of-sight (LOS) link. If the receiver and LED emission source are blocked by obstacles, WiFi can also provide a downlink transmission for a short time to ensure the quality of service (QOS). Thus, the deep integration of visible-light communication and WiFi has been realized and a heterogeneous network is formed.

However, an RF uplink also has some drawbacks. For example, an RF uplink will interfere with the existing system, and occupy the already-crowded spectrum resources. In addition, electromagnetic radiation results when a radio frequency propagates upward. This electromagnetic radiation cannot be allowed in electromagnetic-sensitive environments, such as gas stations, hospitals, and ammunition warehouses. It also greatly reduces the confidentiality of visible-light communication.

At present, the design scheme of an indoor VLC-WiFi network usually depends on the VLC deployment. When planning and deploying indoor lights, the actual needs of the indoor-light users must be fully considered when the complete house design begins. Moreover, the indoor light intensity must meet the international lighting standard of 300–1500 lx [58]. Otherwise, the indoor light is too strong or too weak and may have an impact on human-eye health.

The simple model of the VLC system with the mixed WiFi and VLC uplink is shown in [59] Fig. 38.9. The user's uplink is taken care of by the WiFi-access or APD-access end, and the user's downlink is taken care of by VLC. The ellipses show the communication range of each LED, and the communication range of the WiFi and VLC covers two LED lights. The WiFi access end, APD access end, and VLC end are connected to the external network through wired mode. The WiFi access end

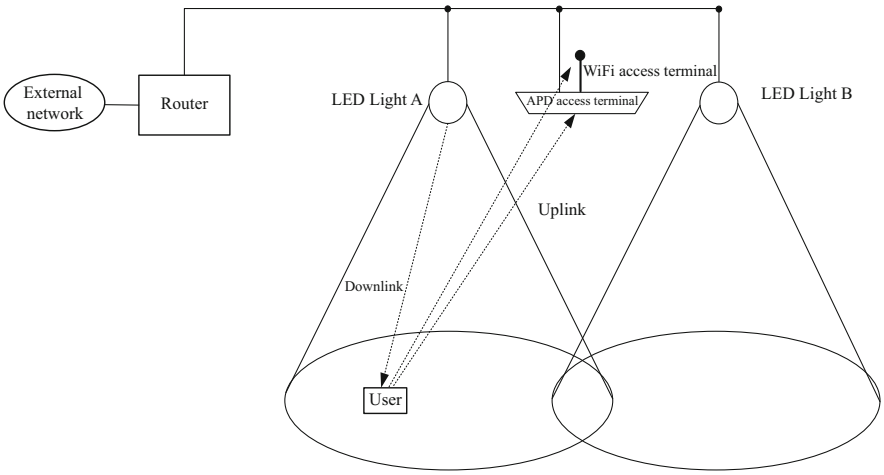


Fig. 38.9 WiFi–VLC system model diagram

and the infrared access end broadcast their respective loads and other relevant information to the VLC end in wired mode.

Power-Carrier Uplinks

Power-line communication (PLC) technology refers to using medium- and low-voltage power lines to transmit data, voice, and other signals. Its advantage is that the coverage is large, the first line is dual-use, and all types of electrical appliances can directly act as network terminals. PLC for internal home networking has mature standards and products and provides a strong competitive solution to solve the "last 300-m/100-m" problem [60].

Power-line communication uses a power line as a channel. Using the principles of spectrum shift and frequency-band division, after one or several modulations, the original signal is moved to a specific frequency band, which is conducive to its transmission in the channel. That is, the signal needing to be sent by the sending end is modulated by power-line modulation technology and then passes through the coupler so that the high-frequency information carried in the signal can be sent in the power line.

After the signal is transmitted to the receiving end, the modulated signal is removed by the coupler and demodulated to obtained the required communication signal. Power-line connections can directly use existing 220 V or 380 V low-voltage power lines as a communication medium to meet the needs of residential communication, but they also can be used in [60] family, office, and other local area networks.

The natural integration of PLC and VLC makes it more convenient for indoor visible-light communication to access the Internet. A PLC–VLC fusion system has good performance and does not require additional wiring; hence, it is widely used in power systems (including smart grids), Internet of Things, and digital homes [61]. However, when the indoor visible light communication system uses an LED as a wireless transmitter, it is necessary to eliminate the problem of information island [62].

PLC–VLC-Fusion Communication-System Structure

At present, a PLC–VLC fusion has two main fusion modes. First, the signal loaded on the power line is demodulated and amplified and then loaded on the LED; this is the decode-forward amplify-forward (DF/AF) mode [63], as shown in Fig. 38.10. In the figure, DF demodulates the signal transmitted on the power line, while AF amplifies the signal transmitted on the power line.

The second mode is the deep fusion of PLC and VLC; that is, the PLC signal is directly loaded onto the LED, without modulation, decoding, or amplification, as shown in Fig. 38.11. An APD is used in the receiving end of both fusion systems to convert optical signals into electrical signals [64].

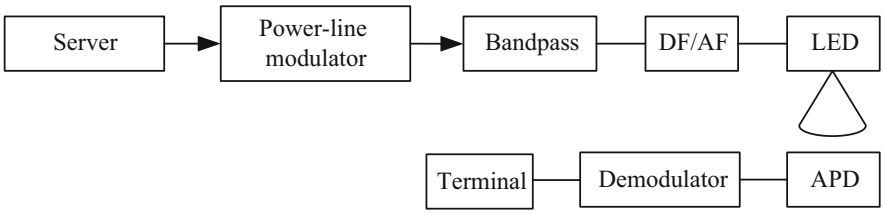


Fig. 38.10 DF/AF fusion method

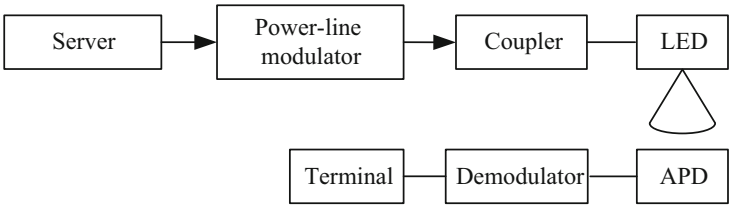
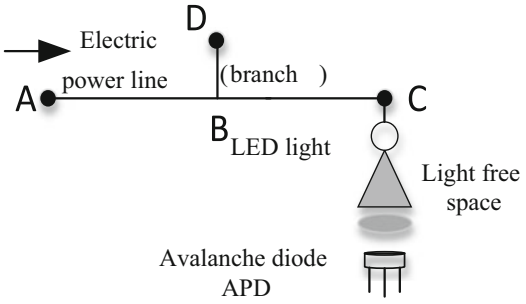


Fig. 38.11 Deep-fusion method

Fig. 38.12 Channel structure model of a PLC–VLC fusion system



Channel Model of a PLC–VLC-Fusion Communication System

Figure 38.12 shows the channel model [65] of the fusion system. The system mainly includes PLC, LED-sending, free-space, and APD-receiving parts. The transmission process is as follows. The signal is first loaded onto the power line, and transmitted to the LED through the power line. Then, the LED converts the electrical signal on the power line to an optical signal and then transmits it. Finally, the APD restores the optical signal to an electrical signal during reception, and the communication is complete.

PLC–VLC-Fusion Communication-System Scheme

At present, there are three PLC–VLC fusion system schemes [66], as follows.

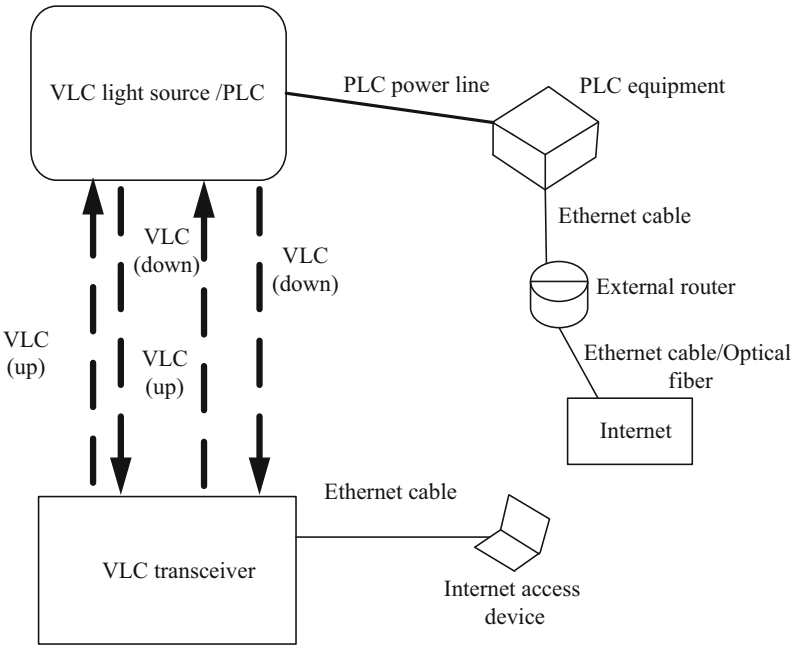


Fig. 38.13 Scheme 1: VLC fixed-position transmission schematic diagram

Scheme 1: The VLC is in a fixed position. Specifically, a special VLC transceiver is arranged below the VLC light source to enable the VLC transmission to be completed from a fixed position. As shown in Fig. 38.13, any Internet access device is connected to the VLC transceiver through an Ethernet cable, and VLC upstream and downstream communication is completed between the VLC light source and VLC transceiver. The VLC light source and external network port device are connected through the PLC device.

The advantage of this plan is that the Internet equipment does not need to be modified; it only needs to access a network line to be connected to the Internet. The earliest VLC and PLC combination scheme is the simplest, and requires the fewest equipment changes [67].

Scheme 2: VLC fixed-point transmission. In this case, multiple VLC devices are equipped with VLC transceivers and can communicate simultaneously with the VLC light. Figure 38.14 shows that several VLC devices can share a VLC light source in the same room. The scheme can also be generalized to have multiple VLC light sources in a room. The VLC light sources can be networked and communicate through a PLC power line, and users can use the nearest VLC light source for Internet communication.

Scheme 3: VLC down, PLC up. As shown in Fig. 38.15, indoor VLC equipment can be uplinked back through the PLC, which effectively reduces the optical-signal interference caused by the light emitted between VLC devices.

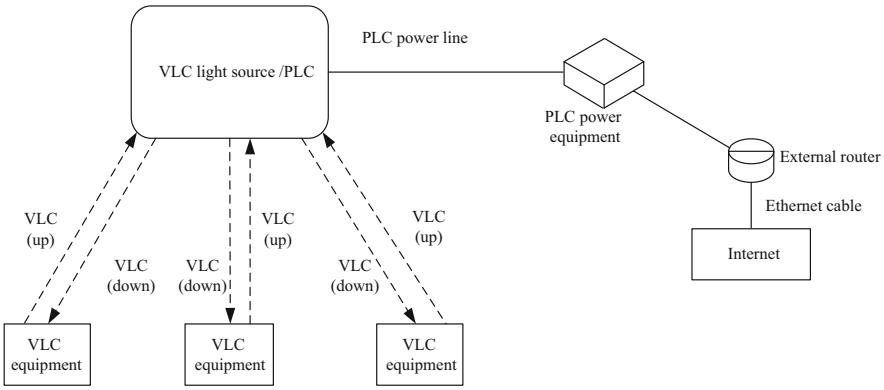


Fig. 38.14 Scheme 2: VLC fixed-point transmission

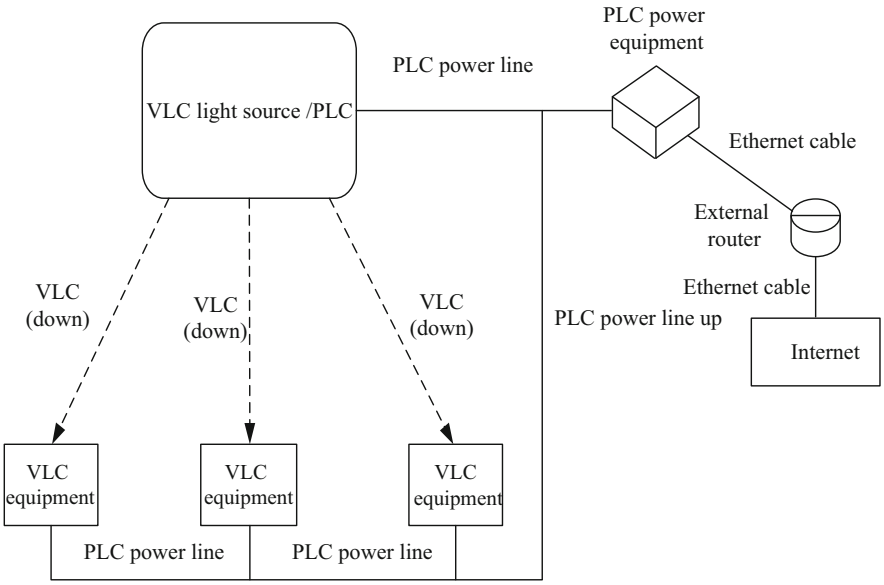


Fig. 38.15 Scheme 3: VLC downlink, PLC uplink

Uplinks with Reverse-Modulation Technology

Downlink application technology tends to be mature; however, the uplink problem has always plagued bidirectional communication systems. The existing visible-uplink solutions are mainly RF, infrared light, and visible light. Other solutions contain drawbacks. For example, the RF-uplink confidentiality is poor and it cannot

be applied in electromagnetic-sensitive environments, such as gas stations, ammunition depots, and nacelle.

Most visible-light and infrared-light uplink structures are more complex. Moreover, the mobile light source carried by the mobile terminal brings more interference, causes discomfort to human eyes, and other problems. Applying reverse-modulation technology to a visible uplink removes the need for an additional light source, as well as the need for a spectrum license. It also ensures human-eye safety, has other inherent advantages, and greatly reduces the complexity of the system structure [68].

Commonly used reverse modulators (modulating retro-reflectors (MRRs) [8]) are angle reflector and cat's eye MRRs. The angle-reflector MRR includes an angle reflector and light modulator. The transmission-type light modulator arranges the angle reflector into the pupil position to change the reflected-light signal strength to achieve reverse modulation. A cat's-eye MRR is composed of a focusing lens and modulating reflector. The modulating reflector is placed in the focal plane of the cat's-eye reverse reflector, and uses reflector-reflectivity or reflector-defocusing changes to control the intensity of the reflected light, to achieve reverse modulation [69].

Structure of a Reverse-Modulation System

The indoor visible-light single-source full-duplex communication system in Fig. 38.16 consists of two parts: the active end and the reverse end. The active end of the system is generally fixed to the ceiling and is responsible for sending information while lighting the room. The reverse end is generally arranged on various types of terminals, and is responsible for providing return data after receiving the downlink information transmitted by the sending end.

The active terminal consists of an Ethernet card, field-programmable gate array (FPGA), analog-to-digital convertor (ADC), digital-to-analog convertor (DAC), LED driver circuit, photoelectric receiving circuit, and signal-processing circuit.

The reverse end includes a defocusing cat's-eye reverse modulator and drive circuit, photoelectric receiving circuit, ADC/DAC circuit, FPGA master module, Ethernet card, etc.

The current shortcoming of reverse-modulation technology applied in the uplink is that the transmission rate is low; however, this technology has the advantages of a simple system, no harm to the human body, no need for a spectrum license, etc. Consequently, reverse modulation has great development potential in the future.

Reverse-Modulation Uplinks

An uplink is a reverse communication link from the reverse end to the active end. Figure 38.17 shows the structure diagram of a reverse communication system. The reverse communication link takes the cat's-eye reverse-modulation system as the

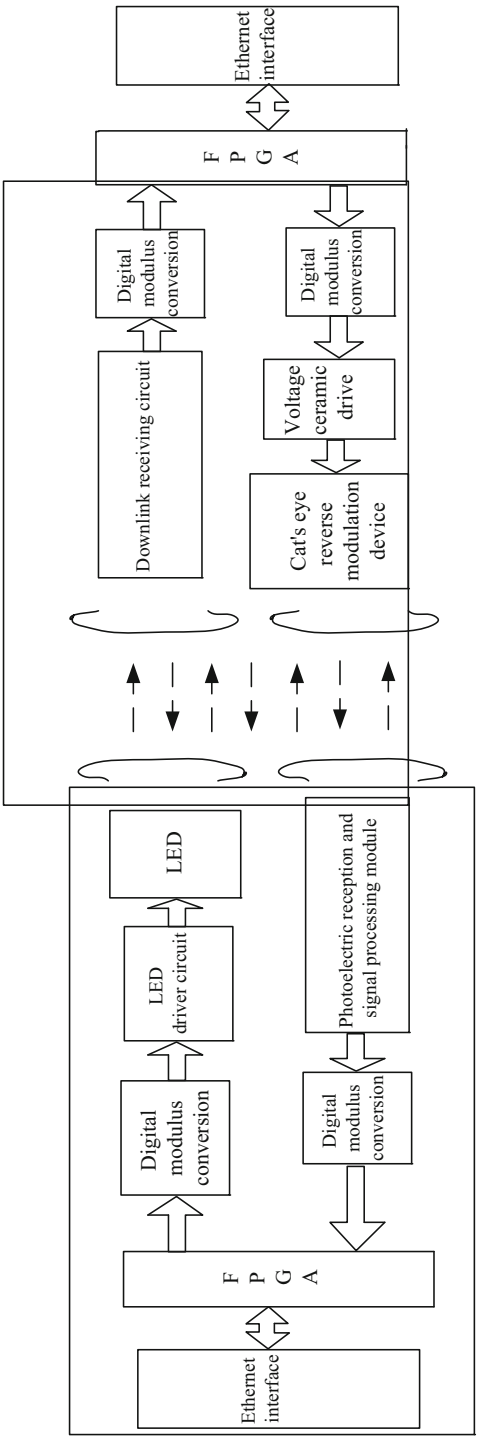


Fig. 38.16 Overall structure of the single-source full-duplex communication system

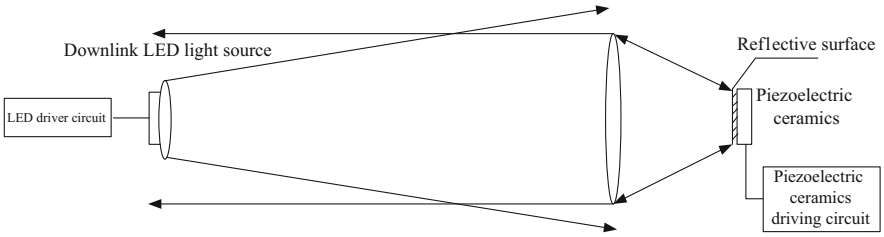


Fig. 38.17 Structure diagram of a reverse-communication system

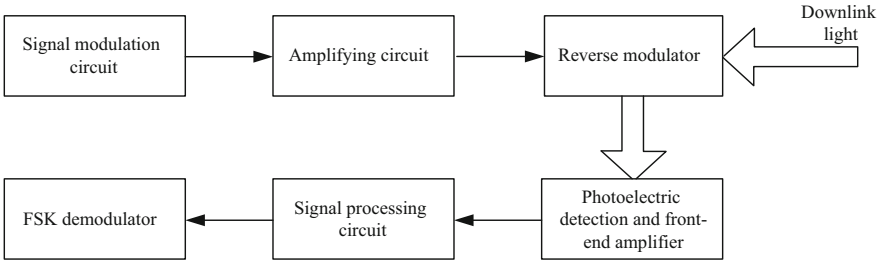


Fig. 38.18 Schematic diagram of the uplink

core. The defocusing cat's-eye reverse modulator includes an optical lens, piezoelectric-ceramic plate, modulation drive circuit, echo receiving circuit, etc.

During the working process, the active end must transmit the modulated light signal to the reverse-modulation end. Owing to the large divergence angle of a white LED beam, the focusing lens of the reverse end and the cat's-eye reverse modulator must be arranged side by side. They are both located in the effective illumination range of the LED light source.

In the reverse transmission, the downlink light signal first goes through the secondary modulation of the reverse modulator, modulates to the downlink light path, and reaches the active end through reflection. At the active end, the optical receiver receives the back optical signal, and after demodulation and processing, the original data information is obtained to complete the duplex communication.

The uplink hardware [68] is composed of a cat's-eye reverse modulator, drive circuit, and signal-receiving circuit. Figure 38.18 shows the block diagram of the uplink.

Summary

This chapter introduced the advantages and disadvantages of existing uplink schemes, as shown in Table 38.4.

Table 38.4 Advantages and disadvantages of common visible-optical communication uplink schemes

Type	RF	Visible light	Reverse modulation	Infrared light	Power carrier
Advantages	Mature system, communication links can be masked, and large channel capacity	Simple system structure, large communication capacity, and no need to design an uplink system	System structure, no need for a spectrum license, ensures the safety of human eyes	Mature system, simple structure, single, low system cost	Long transmission distance, strong anti-interference ability, high infrastructure utilization rate
Disadvantages	Incompatible with existing systems, RF interference exists, and occupies spectrum resources	Uplink light is visible to the human eye, has a visual impact, and limits the applications	Low transmission rate	Short communication distance, low communication bandwidth, the human eye is sensitive to infrared light, and poor security	Low transmission rate, low security

The prospects of visible uplinks are very broad, and further research in the future includes the following:

1. Researchers should investigate the matching degree of uplink and downlink transmission rates and improve uplink transmission rates and distances to meet the needs of different scenarios.
2. Researchers should investigate combinations of uplinks and other technologies, such as a heterogeneous network system combining visible light as the downlink and WiFi as the uplink, and full-duplex communication systems in which both uplinks and downlinks use VLC to improve the uplink-transmission performance and reliability.
3. Researchers should consider the application of reverse-modulation technology in the uplink scheme, because it ensures the security of the human eye and simplifies the system structure.
4. Researchers should apply uplink technology to life, such as Internet of Things, Internet of vehicles, drones, industrial automation, smart homes, mobile communication, and other fields.

References

1. Sun Dazhuang. Research on LED-based visible light bi-directional communication technology [D].: Nanjing University of Aeronautics and Astronautics, 2021.
2. Zhang Huan. Design of visible light full duplex communication system [D]. Xi'an: Xi'an University of Technology, 2018.

3. Ayyash M, Elgala H, Khreishah A, et al. Coexistence of WiFi and LiFi toward 5G: concepts, opportunities and challenges[J]. IEEE Communications Magazine, 2016, 54(2): 64–71.
4. Kahn J M, Barry J R. Wireless infrared communications[J]. Proceedings of the IEEE, 1997, 85(2):265–298.
5. Wu Menglong., Guo Jia., Liu Wenkai., et al. Scheme and implementation of bi-directional visible light communication system with LED as transceiver element[J]. Journal of Optics, 2018, 38(4): 4–8.
6. Cossu G, Corsini R, Khalid A M, et al. Bi-directional 400 Mbit/s LED-based Optical Wireless communication for Non directed Line of Sight Transmission[J]. IEEE, 2014, 22(10): 1–3.
7. Gfeller F R, Bapst U. Wireless in-house data communication via difuse infrared ra-diation[J]. Information Systems, 1980, 5(3):1474–1486.
8. Ke Xizheng, Zhang Huan. Indoor visible uplink communication system based on retrograde modulation[J]. Advances in Lasers and Optoelectronics, 2018, 55(02):115–122.
9. Qiu Hao, Wang Jinyuan, Xu Zhiyong, et al. Research on optical communication technology with single-source full-duplex reverse modulation[J]. Optoelectronics and Lasers, 2015, 26(12): 2314–2319.
10. Ding Kun., Zhang Tao, Zheng Weibo. Portable optical antenna device based on visible light communication technology[J]. Optical Communication Research, 2017, 199(01):76–78.
11. Liu YF, Yeh CH, Chow CW, et al. Demonstration of bi-directional LED visible light communication using TDD traffic with mitigation of reflection interference[J]. Opt Express. 2012, 20 (21):19–24.
12. Xu Shanhe, Xiao Shali. A novel duplex communication system with a single light source [J]. Semiconductor optoelectronics, 2015, 36(03):447–450.
13. Rahaim M B, Vegni A M, Little T D C. A hybrid radio frequency and broadcast visible light communication system[J]. IEEE, 2011, 18(52):792–796.
14. Basnayaka D A, Haas H. Design and Analysis of a Hybrid Radio Frequency and Visible Light Communication System[J]. IEEE Transactions on Communications, 2017, 65(10):4334–4347.
15. Lu, Tingwei, Wang, Zeping., Liu, Meng., et al. Design of a full-duplex Ethernet communication system based on visible light communication technology[J]. Journal of Electronics, 2022, 50(01):45–53.
16. Schmid S, Corbellini G, Mangold S, et al. LED-to-LED visible light communication networks [J]. ACM, 2013, 35(06): 1–10.
17. Zheng Zong, Liu Lu, Hu Weiwei. Uplink scheme for visible light communication[J]. ZTE Communication Technology, 2014, 20(6):8–11.
18. Gfeller F R, Bapst U. Wireless in-house data communication via diffuse infrared radiation [J]. Proceedings of the IEEE, 1979, 67(11): 1474–1486.
19. Zhang S, Zhao L, Song S, et al. Full-duplex transmission without an uplink light source for a millimeter-wave radio over a free-space optical system[J]. Appl Opt. 2022, 61 (28):8323–8339.
20. Komine T, Lee J H, Haruyama S, et al. Adaptive equalization for indoor visible-light wireless communication systems[J]. IEEE, 2005: 20(09):294–298.
21. Ning W, Yang Q, Wuqu W, et al. Visible light communication based inellieinttraffic light system: designing and implementation[J]. IEEE Access, 2018, 2(17):4–15.
22. Karem Rana, Ahmed Mehaseb, Newagy Fatma. Resource Allocation in Uplink NOMA-IoT Based UAV for URLLC Applications[J]. Sensors, 2022, 22(4):5–44.
23. Olejniczak Alicja, Błaszkiewicz Olga, Cwalina Krzysztof K. Software-Defined NB-IoT Uplink Framework-The Design Implementation and Use Cases[J]. Sensors, 2021, 21(24):8–16.
24. Nakamura Mio, Nishiyama Hiroki. Potential of Cooperative SIC for Uplink NOMA in Multi-Cell Network[J]. IEICE Communications Express, 2022, 11(8):497–502.
25. Kim Hyerin, So Jungmin. Improving Spatial Reuse of Wireless LAN Uplink Using BSS Color and Proximity Information[J]. Applied Sciences, 2021, 11(22):154–160.
26. Rosenthal James, Sharma Apoorva, Kampianakis Eleftherios. A 25 Mbps, 12.4 pJ/b DQPSK Backscatter Data Uplink for the NeuroDisc Brain-Computer Interface[J]. IEEE transactions on biomedical circuits and systems, 2019, 13(5):858–867.

27. Cahyadi WA, Chung YH. Experimental demonstration of indoor uplink near-infrared LED camera communication. [J]. Opt Express. 2018; 26(15):19657–19664.
28. Wang K, Kandeepan S. Experimental demonstration of an indoor optical wireless communication system with a waveform index modulated uplink. [J]. Opt Lett. 2022, 47(13):3187–3190.
29. Yun Wenyue. Research on modulation and demodulation technology of indoor visible light communication system[D]. Changchun, Jilin, China: Changchun University of Technology, 2008.
30. Kang Ye, Ke Xizheng. Multidimensional coding in visible optical communication[J]. China Laser, 2015, 42(2): 130–149.
31. Song J, Ding W, Yang F, et al. An indoor broadband broadcasting system based on PLC and VLC[J]. IEEE Transactions on Broadcasting, 2015, 61(2):299–308.
32. Chen Q. R., Zhang T., Zheng W. B., et al. Research status and application prospects of white LED-based visible light communication[J]. Semiconductor optoelectronics, 2016, 37(4): 455–460.
33. Xingzheng Ke, Ke Dong. Optical Wireless Communication[M], Springer Press, 2022:77–83, 127–140.
34. Zhang Peng, Wang Tianshu, Yang Guowei, et al. Performance evaluation of full duplex reverse modulation reply space optical communication system[J]. Infrared and Laser Engineering, 2015, 44(08):2506–2510.
35. Duan Zhiwen. Spectral efficiency enhancement of uplink in de-cellular large-scale MIMO system[D]. Guilin, Guangxi: Guilin University of Electronic Science and Technology, 2022.
36. Liu Yang, Cai Xiping, Lin Lang, et al. Research on interfacing technology of LED visible light communication network with Ethernet [J]. Optical Communication Technology, 2019, 43(3): 1–4.
37. Sun Xinxin, Wang Yanjun. Design of Ethernet communication system based on visible light communication technology[J]. Information Recording Materials, 2022, 23(10):170–172.
38. Feng Qing, Fang Anan, Wang Rongrong. Analysis of improved indoor visible wireless LAN based on OCDMA[J]. Television Technology, 2014, 38(03):384–391.
39. Zeng Yu, Li Wenjie, Zhang Guanghui, et al. Research on visible optical communication and access technology[J]. Telecommunications Technology, 2016(05):71–75.
40. Zhu Mangjing, Wang Yuhao, Liu Xiaodong. Research on visible optical communication relaying technology[J]. Optical Communication Research, 2021(03):64–71.
41. Wang X, Wang C, Wang C, et al. Research on uplink scheduling algorithm for LTE system [J]. Communication Technology, 2020, 53(03):667–672.
42. Jingmiao Wu. Research on uplink performance improvement in heterogeneous networks [D]. Inner Mongolia: Inner Mongolia University, 2019.
43. Komine T, Nakagawa M. Integrated system of white LED visible-light communication and power-line communicationL[J].IEEE Transactions on Consumer Electronics, 2003, 49(1): 71–79
44. Amirshahi P, Kavehrad M. Broadband access over medium and low voltage powerlines and use of white light emittingdiodes for indoor communications[J]. IEEE, 2006, 35(09):897–901
45. Rajesh A, Nakkeeran R. Performance analysis of integrated system under impulse noise and multipath channel using Tur-bo coded OFDM[J]. IEEE Transactions on Consumer Electronics, 2008, 21(07):254–259.
46. Alavi S E, Supa'at A S M, Idrus S M. Integrated system of visible free space optic with PLC [A]. IEEE. Proc of 2009 IEEE 9th Malaysia International Conference on Communications [C]. Kuala Lumpur, Malaysia: IEEE, 2009:271–275.
47. Ding W, Yang F, Yang H, et al. A hybrid power line andvisible light communication system for indoor hospital appli-cations[J]. Computers in Industry, 2015, 6(8):170–178.
48. Song J, Ding W, Yang F, et al. An indoor broadband broad-casting system based on PLC and VLC[J]. IEEE Transactions on Broadcasting, 2015, 61(2):299–308.
49. Yi Fan. Research on uplink multi-user transmission technology for visible optical communication system[D]. Southeast University, 2017.

50. Chi, Nan. Key technologies for high-speed visible light communication [M]. Beijing: People's Post and Telecommunications Publishing House, 2019:22–24.
51. Bergh A A, Copeland J A. Optical sources for fiber transmission systems[J]. Proceedings of the IEEE, 1980, 68(10): 1240–1247.
52. Saengudomlert P. Transmit beamforming for line-of-sight MIMO VLC with IM/DD under illumination constraints[J]. IEEE, 2015:1–4.
53. Yang, Yingping., Hu Cangkui., Hu, J. H.. Optoelectronic technology [M]. Beijing: Machinery Industry Press, 2014:97–98.
54. Diamantoulakis P D, Karagiannidis G K, Ding Z. Simultaneous lightwave information and power transfer[J]. IEEE Transactions on Green Communications and Networking, 2018, 2(3): 764–773.
55. Obeed M, Dahrouj H, Salhab A M, et al. DC-bias and power allocation in cooperative VLC networks for joint information and energy transfer[J]. IEEE Transactions on Wireless Communications, 2019, 18(12): 5486–5499.
56. Mohanad O, Anas M. S., Salam A. Z., et al. Joint optimization of power allocation and load balancing for hybrid VLC/RF networks[J]. IEEE/OSA Journal of Optical Communications and Networking, 2019, 10(5):553–562.
57. Z. Du, C. Wang, Y. Sun, et al. Context-aware indoor VLC/RF heterogeneous network selection: reinforcement learning with knowledge transfer[J]. IEEE Access, 2018, 6(07):33275–33284.
58. IEEE Standard for Local and Metropolitan Area Networks-Part 15.7:Short-Range Optical Wireless Communication Using Visible Light, IEEE Std 802.15.7-2011, Sep. 2011, pp. 1–309.
59. Gao Yang. Research on WiFi and infrared based hybrid link technology in indoor visible light communication system[D]. Xi'an University of Electronic Science and Technology, 2016.
60. Yang Liwei, Hou Cong, Cheng Qiang. Research on visible optical communication and power line convergence technology[J]. Telecommunications Network Technology, 2014(02):58–61.
61. Lin Y J, Latchman H A, Lee M, et al. A power line com-munication network infrastructure for the smart home[J].IEEE Wireless Communications, 2008, 9(06):104–111.
62. Ma Hao, Lampe L, Hranilovic S. Integration of indoor visible light and power line communication systems[J] IEEE, 2013, 11(10):291–296.
63. Chi Nan. LED visible light communication technology [M]. Beijing: Tsinghua University Press,2013.
64. Cai Longfei. Research on converged channels for power line and optical communication [J]. Wireless Interconnection Technology, 2020, 17(21):3–4+11.
65. Gao Jue, Zhang Jian, Yang Hui. Reference model of converged channel for indoor power line communication and visible light communication[J]. Computer Engineering and Science, 2018, 40(06):1062–1066.
66. Wang Yi, Xu Zhenzhao, Wang Nan. The combination of power line communication and visible light communication and its mobile application[J]. Telecommunications Express, 2019(11): 5–10.
67. Xu Jidong, Wang Lin, Yang Liwei, et al. Research report on the combination technology of visible optical communication and power line[R]. Beijing: China Communications Standardization Association, 2014.
68. Sun Huauan., Zhang Laixian., Zhao Y. Z., et al. Advances in reverse modulated free-space laser communication technology[J]. Advances in Laser and Optoelectronics, 2013, 50(4):56–63.
69. Yu Feng Yang, Mingzheng Jiang, Ying Zhang, et al. Design of a full-duplex visible light communication system based on a single light source[J]. Advances in Lasers and Optoelectronics, 2019, 56(01):78–86

Research Progress on Indoor Visible Light-Source Layouts

39

Contents

Introduction 1464
Visible-Light Communication Systems 1464
Principle of Visible-Light Communication Systems 1464
Channel Model and Light-Source Layout for Visible-Light Communications 1466
Domestic and Foreign Research Progress on VLC Light-Source Layouts 1467
Geometric-Layout Methods 1467
Light-Source-Location Optimization Methods Based on Intelligent-Optimization Algorithms 1469
Methods for Optimizing the Light-Source Parameters Based on Intelligent Optimization Algorithms 1472
Summary of Research Progress 1476
Visible-Light Communication Light-Source-Layout Analysis 1476
Light-Source-Layout Principles 1476
Light-Source Spatial Layout 1477
LED-Array Light-Source Layouts 1481
Key Algorithms for Light-Source Layout Parameters 1481
Research on VLC Light-Source Layouts at Xi'an University of Technology 1484
Received-Power Analysis 1484
Optimal Design of a Light-Source Layout 1485
Layout-Model Simulation Analysis 1486
Summary 1489
References 1491

Abstract

In the indoor visible light communication system, in order to ensure that many users in different directions can get better communication quality and achieve uniform optical power coverage on the receiving surface, it is necessary to reasonably optimize the layout of the light source. Reasonable light source layout design can achieve stable signal transmission, improve the signal-to-noise ratio of the transmitted signal and reduce inter-symbol interference, and help to improve the communication rate. This chapter briefly describes the proposal, development, and

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_39

unique advantages of visible light communication, summarizes the principle and channel model of visible light communication system, summarizes and compares the light source layout and design of visible light communication system, explains its development context and research progress, and finally summarizes the development direction of visible light communication layout and the development trend of design methods.

Keywords

Visible light communication · Channel characteristics · Light source layout · Receiving power

Introduction

At present, the scarcity of wireless-spectrum resources requires a more efficient and secure communication method. Optical-communication technology is gradually replacing traditional wireless-communication methods, owing to its high signal-to-noise ratio, fast modulation rate, good confidentiality, and other advantages. Visible-light communication (VLC) has become a hot spot. Visible-light communication technology can transmit information while lighting a room and solves the problems of broadband access and high-speed data transmissions [1–6].

Most existing VLC technologies use white light-emitting diodes (LEDs) as the signal source. In an indoor environment, each LED can be used as an optical access point, and multiple LEDs in a room can constitute an optical cellular wireless network [7]. Because the coverage of a single LED is limited and cannot meet the lighting and communication requirements of all users in a room, the LED layout becomes very important in VLC systems. A reasonable layout can distribute the light sources evenly, improve the signal-to-noise ratio of the transmitted signal, reduce inter-code interference, and greatly help to improve the data-communication rate [8].

Visible-Light Communication Systems

Principle of Visible-Light Communication Systems

In indoor VLC systems, visible light is used to transmit information. LEDs, which are commonly used in daily life, are often used as light sources [9]. Visible-light communication systems use light that is visible to the human eye, with wavelengths between 380 nm and 780 nm, for data transmission [10, 11], as shown in Fig. 39.1.

A visible-light communication (VLC) system consists of a modulation module, driver circuit, LEDs, channel, optoelectronic conversion module, circuit amplifier, equalizer, judgment circuit, and other modules.

The signal is modulated and converted into a form suitable for transmission to the driver circuit. The role of the driver circuit is to perform an electrical/optical

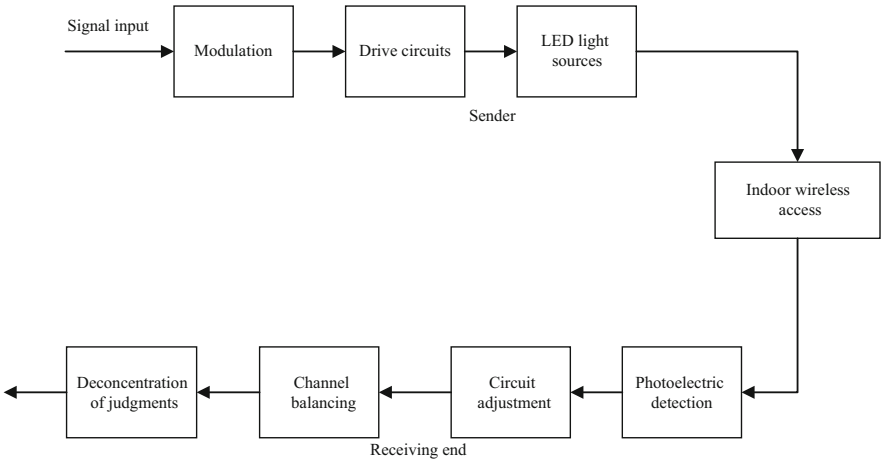


Fig. 39.1 Block diagram of an indoor visible-light communication system [12]

conversion on the modulated signal, provide the appropriate drive-current size, and load it onto the LED light source. The light signal is then transmitted through an indoor wireless channel. The photodetector receives the visible-light signal from the light source and converts the light signal into an electrical signal under illuminated conditions, enabling the information to be recognized and processed.

The VLC-system LED light source radiates illumination following the Lambert model, with radiant energy at a directional angle of

$$I(\phi) = I(0)COS^{m}(\phi) \tag{39.1}$$

where m is the Lambert factor. $I(0)$ is the light intensity when the relative angle between the light source and receiver is 0°, at which point, the total energy of the light source, P_t, can be expressed as

$$P_t = 2\pi I(0)\frac{1}{m+1}. \tag{39.2}$$

The expression for the Lambert factor m in the above equation is

$$m = -\frac{1}{log_{\,2}\,cos\left(\phi_{1/2}\right)} = \frac{-In2}{In\ cos\left(\phi_{1/2}\right)}. \tag{39.3}$$

The optical gain of an ideal nonimage concentrator with an internal refractive index n can be expressed as

$$g(\theta) = \begin{cases} \dfrac{n^2}{\sin^{\,2}(FOV)} & 0 \le \theta \le \text{FOV} \\ 0 & \theta > \text{FOV} \end{cases} \tag{39.4}$$

FOV is the field of view of the receiving device and θ is the light angle of incidence. The strength of the received energy determines the quality of communication in a VLC system. The channel characteristics affect the received field of view and the angle of incidence of the incident light, as well as parameters such as the transmission power and receiving area of the photodiode. To ensure stable visible-light communication, the VLC channel model must be analyzed.

Channel Model and Light-Source Layout for Visible-Light Communications

Indoor visible-light communication systems have LEDs and receivers installed at different positions and angles. Figure 39.2 shows a model of the most significant indoor line-of-sight and non-line-of-sight links.

In an indoor VLC system, signals are transmitted and received through three processes. However, owing to obstructions, such as walls and partitions in the room, the transmitter transmits the signal to the receiver through two different channels.

In the first channel, called a direct-vision link, the optical signal is sent directly from the transmitter to the receiver, without being obstructed by obstacles. The advantages of a direct-vision link are that it has strong anti-interference performance, its bandwidth is relatively high, and it can support high-speed data transmissions. However, a direct-vision link limits the locations of the transmitter and receiver.

Some signals cannot reach the receiver directly from the transmitter, but are obstructed by obstacles during transmission. This type of channel is called a

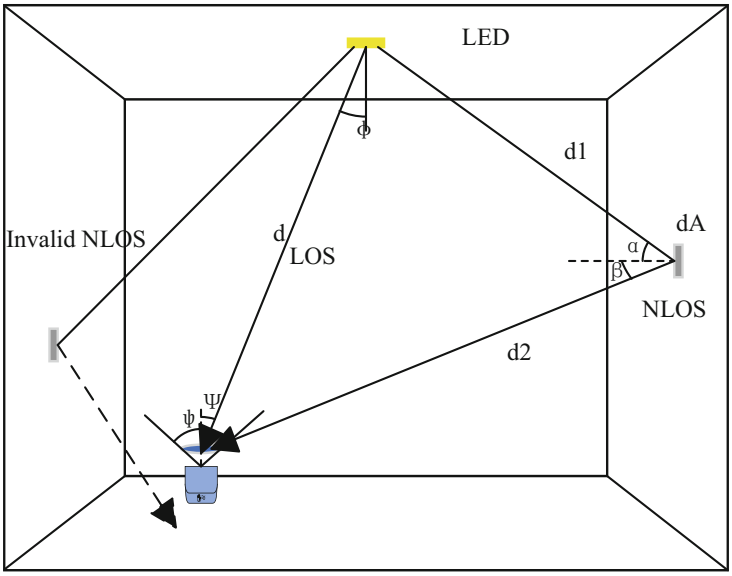


Fig. 39.2 Indoor VLC channel model [13]

non-direct vision link. Owing to the poor transmission path of a non-direct vision link, the receiver must receive direct light from one of the LED light sources to reduce the communication and illumination blind spots [14].

Therefore, the light-source layout of indoor VLC systems should be optimized, to achieve the necessary illuminance distribution and ensure the required receiving-power distribution. The layouts of VLC-system light sources can be divided into indoor-communication layouts, vehicle-communication layouts, and customized layouts for special occasions with special needs, according to the actual requirements. They can also be divided into LED lamp-bead layouts and relative light-source position layouts, according to the research objective.

This chapter summarizes the layout and design methods for VLC-system light sources and compares different research results to illustrate future research directions and development trends.

Domestic and Foreign Research Progress on VLC Light-Source Layouts

Researchers have optimized and improved the layouts of light sources in VLC systems such that the optimized layout meets both the light-illumination and communication-quality criteria. The broad improvement options include geometric layouts, and using intelligent optimization algorithms for light-source location and parameter optimization.

Geometric-Layout Methods

In 2004, Fan et al. [15] designed a square light-source layout in a 5 m × 5 m × 3 m room. They analyzed the differences in the light paths of multiple LEDs using different layout methods and investigated the differences between direct and non-direct links in the room. The effects of the LED layout on the signal-to-noise ratio (SNR) and transmission rate of the VLC system were investigated. Although this layout improved the illumination of the indoor light source, it lost a large amount of energy from the light source. Simple position optimization is not sufficient for stable communication; therefore, researchers have introduced multi-lamp-source models.

In 2006, Sivabalan et al. [16] analyzed the effects of multiple LED light sources on the power distribution and channel bandwidth. The results showed that a uniform light-intensity distribution was obtained when multiple light sources were used in a room. The key to this was the locations of the light sources. The optimal light-source locations were symmetrically distributed throughout the room, with respect to the floor and ceiling.

In 2007, Ding and Ke from Xi'an University of Technology [17] proposed a model for the layout and transmission of light sources for indoor optical communication. They analyzed the relationship between the layout and received power

distribution of the system, calculated the DC gain, and arrived at the optimal layout for using four LED light sources in a room. However, as this was the initial stage of indoor optical-communication research, only the direct-light case was considered, for the convenience of the simulation experiments, which had significant limitations.

In 2010, Nguyen et al. [18] devised a column-array layout to replace the common light-source layout method with LED light sources as the center point. The array had a high offset rate at all locations of the walls in the room to create a blind-spot-free indoor data-transmission range. Although innovative, the communication performance and appropriate number of central LEDs were not investigated in depth.

In 2012, Siwamogsatham et al. proposed a system that transmitted multiple beams simultaneously in a uniform circular-array layout with a transmitter source that finally converged for the output. They created a multiple-input single-output beamforming system with a tunable optical transmitter in a uniform circular-array layout [19] to achieve higher channel capacity and better transmission performance.

In 2013, Chen [20] designed a novel layout model that ensured illumination while reducing the number of LEDs. By focusing on the non-direct-vision link characteristics and the factors that created multipath effects in the system, this study ensured the system reliability while improving the uniformity of the illumination and reception rates.

In 2016, Wang et al. [21] proposed an elliptical distribution layout based on the room aspect ratio in a 5 m $\times$ 4 m $\times$ 3 m indoor space. The LEDs were reasonably arranged according to the room size to prevent performance degradation, owing to coverage blindness in indoor VLC systems. The simulation results showed that the elliptical LED layout better improved the communication quality.

Chen [22] proposed a double-button-line light-source layout for same-size rectangular rooms to solve the problem of uneven illumination and poor communication quality in the system, using the same number of LEDs. However, this optimization was not supported by mathematical theory, and the lower received power in the corners of the room, compared to the center, led to a reduction in communication quality.

In 2017, Zhao and Peng [23] proposed a 4 + 1 layout, based on the traditional layout, and optimized the original layout. The 4 + 1 layout method needed to consider other aspects, in addition to light uniformity and communication effects.

In 2017, Chen and Garg et al. [24, 25] presented quadrilateral and hexagonal LED-layout methods to improve indoor visible-light coverage. However, they only considered the number of LED positions in a specific environment and did not provide the specific positions and layout numbers of the LEDs; thus, this method was not universal. In 2017, Zhao et al. [26] developed a new layout array and constructed a model consisting of a five-LED array. This scheme reduced the energy consumption of the system.

In In 2017, Wang and Che et al. [27] designed a new circular-layout scheme for a square room. This method enabled the communication system to consume less energy and achieve the best communication effect, with a minimum number of LEDs. The scheme proposed in this study overcame the limitations of the original idea and made the design more flexible.

In the same year, Zhao and Song et al. [28] proposed a scheme completely different from the conventional layout. They used a central-compensation scheme consisting of five arrays, where the considered usage scenario was no longer a square room but a narrower indoor space. After the simulation, it was concluded that the center-compensated structure was the best solution to satisfy both the lighting and communication requirements.

In 2018, Zhao et al. [29] used a 4 m × 4 m × 3 m room as a model and established a ring light-source layout with a single LED array and a light strip under the original array-layout model using the light-intensity compensation principle. This paper proposed a scheme for the layout of ring light sources that considered the communication quality while ensuring the uniformity of the light intensity. It provided an effective optimization method for light-source layouts for indoor visible-light communication.

In 2019, Sharma et al. [30] proposed four types of LED patterns in a defined-size indoor room: diagonal, horizontal centerline, horizontal and vertical lines, and finally, two diagonal lines. They concluded that the two-diagonal design provided the maximum emitted power.

In 2019, Chen et al. [31] proposed a square-array plus circle layout, which consisted of four $N \times N$ squares and m LEDs arranged in a circle. Compared to the traditional square-array layout, the optimized layout had 48 fewer lamp sources. Moreover, the uniformity was improved after using the principle of minimum-illumination mean-square difference for power distribution.

In 2020, Jiang et al. [32] proposed a ring-angle-compensated array light-source layout to study the layouts of array light sources with different arrangements. This method effectively improved the illumination, received power, and signal-to-noise ratio at the receiving end. It also enhanced the signal quality, which was beneficial for the implementation of high-speed VLC systems.

In 2021, Wang et al. [33] proposed the minimum mean-squared error principle to design a circular-layout scheme. Based on this, two types of layouts, circular and rectangular, were used to simulate the received-power distribution and signal-to-noise ratio distribution. The received-power distribution of the non-direct-view link outperformed that of the direct-view link when the number of LEDs was the same; however, the reflection from the wall surface reduced the signal-to-noise ratio. This study provided a new concept for light-source layouts for indoor VLC systems and establishing indoor VLC links (Table 39.1).

Light-Source-Location Optimization Methods Based on Intelligent-Optimization Algorithms

In 2009, Cao et al. [34] proposed an improved particle-swarm optimization (PSO) algorithm. The algorithm effectively avoided the problem of easily falling into local extremes and reduced the accuracy of the optimization calculation when the number of particles remained constant. The algorithm could effectively optimize the layout when there were multiple light sources.

Table 39.1 Research progress in geometric-layout methods

Year	Researchers	Study content	Progress
2004	Fan et al.	Square diode layout, and differences in channel characteristics between different layouts	Derived the effect of different light-source layouts on data-transfer rates
2007	Ding	Developed a transmission model for indoor optical communication based on four LEDs. Analyzed the system layout in relation to the received power distribution	Optimal layout of four LED light sources based on a DC-gain design
2010	Nguyen	Used f an array layout with a high offset ratio	Layout enabled data to be transmitted without blind spots
2012	Siwamogsatham	Uniform circular-array layout with multiple beams emitted simultaneously from the source	Achieved a higher channel capacity and more stable transmission performance
2016	Wang	Proposed an elliptical distribution model based on room aspect ratios	Used an oval LED layout for better communication quality
2016	Chen	Proposed a lemniscate layout suitable for rectangular rooms	Lemniscate layout had a higher uniformity of illumination than other layouts, ensuring uniformity of communication and illumination
2017	Chen	Proposed quadrilateral and hexagonal LED layout methods to improve indoor visible-light coverage	Solved the light-coverage problem in specific environments, but was not universal
2017	Wang	Designed a new circular layout in a square room	Enabled communication systems to consume less energy and achieve the best possible communication with the minimum number of LEDs
2018	Zhao	Modeled the layout of a ring light source with single LED arrays and light strips	Provided communication quality with guaranteed light-intensity uniformity
2019	Chen et al.	Proposed a layout model of a square-array plus circle	Illumination requirements of the room were met with an improved mean-square and uniformity
2020	Jiang et al.	Proposed a ring-angle compensated array light-source layout	Improved illumination, received power, and signal-to-noise ratio at the receiving end; also enhanced signal quality

In 2013, Guan et al. proposed a single-LED layout scheme and an array LED-based layout scheme [35] to obtain an optimal light-source layout scheme, and improved the layout scheme using PSO algorithms. The results showed that the proposed LED-layout method based on the PSO algorithm outperformed the random method.

In 2015, Fu [36] used a genetic algorithm to optimize the position of light sources and reasonably designed an adaptation function. The results showed that the illumination uniformity on the receiving surface reached 97.4% after optimization. The uniformity was improved; however, the distribution of light-source positions was too dispersed for practical use.

In 2016, Nicolas K et al. [37] established an analytical model for the optimal layout of light sources using a novel downhole positioning algorithm that improved the accuracy of the position by controlling the number of light sources and providing a reasonable position layout.

In 2016, Gao [38] proposed a light-source layout for a VLC system in a spacious lighting environment. They analyzed the VLC system-design scheme, adjusted the layout while reducing the number of LEDs in the array, and used a combination of simulated annealing and genetic algorithms to optimize the light-source positions to achieve uniform lighting and communication stability. However, this algorithm was based on the overall optimization of the position, making the searched position highly random. The light-source distribution was not aesthetically pleasing and had low practicality.

In 2016, Ajmani et al. [39] proposed a novel algorithm that estimated the position of a receiver relative to a reference light source when the receiver received signals from multiple LEDs. The optimal beam radius between two LEDs in this array was derived by analyzing parameters, such as the room size, number of light sources, and spacing.

In 2017, Liu and Wang [40] proposed the use of a genetic algorithm for gene density to arrange the layout of a specific room and obtain an ideal power distribution by optimizing the layout of the light-source locations. This algorithm calculated the light-source distribution scheme by itself; therefore, it could be very effective in avoiding manual errors or ideological limitations.

In 2018, Liu [41] proposed a novel light-source layout scheme to improve the signal uniformity of indoor visible-light communication. The method was optimized using a hybrid particle-swarm algorithm to improve the communication performance of the system.

In 2019, Zhai et al. [42] proposed an LED lighting-layout optimization scheme using an artificial fish-swarm algorithm to investigate two different light-source layout methods: rectangular and elliptical. By minimizing the power variation of the receiving surface as the objective function and optimizing the position of the LEDs, an optimal LED-layout pattern was obtained. Further improvements were made to the power coefficient of the LEDs.

In 2020, Liu et al. [43] proposed a light-source layout model based on a multiple-swarm genetic algorithm, which was developed based on a traditional genetic algorithm to find the optimal LED light-source location information. The results showed that the optimized light uniformity reached 91.73% and the illuminance range was 889–1009 lx. Both of these were significantly better than those based on the traditional layout model under a genetic algorithm. Thus, this provided a new reference method for LED light-source layouts.

Ma et al. [44] used particle-swarm algorithms to satisfy the overall light-intensity lighting conditions in the light-source layout in a cabin-lighting communication

system. They ensured that the light power met the requirements, and optimized the position and angle of each light source to minimize light-intensity fluctuations.

In 2021, Li [45] improved the LED layout to address the problems of user distribution and inter-user interference in indoor VLC systems. They used white LEDs as lighting and information transmitters to meet the user-experience quality and demand. To address the distribution of users in different cases, they proposed an algorithm, based on the quality of user experience, to improve the LED layout.

In the same year, Li et al. [46] proposed an improved artificial bee-colony algorithm to optimize the position and layout of light sources to find their specific positions for an optimal lighting effect. They compared the lighting effect of rooms with different specifications. The results showed that the standard lighting-layout requirements could be met without the general light-source layout for irregular rooms to achieve lighting effects and save energy.

In 2022, Qu [47] proposed an evolutionary-algorithm scheme for maximizing the LED layout, based on signal-to-noise coverage. They simulated the illuminance and signal-to-noise ratio of polygonal and circular layouts in a 5 m $\times$ 4 m $\times$ 3 m room, and continuously improved the light-source coordinates to find an optimized layout. The effect of the orientation of the LED arrangement on the performance of the room system was investigated, and this study provided a new method of considering the layout of indoor light sources (Table 39.2).

The above light-source-layout optimization methods all optimize for the location of the light source. Although the light uniformity of the receiving surface is improved and lighting standards are met, to a certain extent, optimizing the light-source location as a whole through intelligent-optimization algorithms is not sufficiently comprehensive. This leads to very random locations, a less aesthetically pleasing distribution of light sources, and lower practicality.

The light-source parameters can be further optimized or the position and parameters can be co-optimized to achieve uniform lighting and improve the communication quality. The following describes light-source parameter-optimization methods based on intelligent optimization algorithms.

Methods for Optimizing the Light-Source Parameters Based on Intelligent Optimization Algorithms

In 2012, Ding et al. [48] proposed a new method for fully diffusing indoor VLC using a custom-evolved algorithm to dynamically optimize the signal-to-noise ratio over the communication layer in multiple scenes. They controlled the optical power of individual LED-lighting devices to achieve fully diffuse indoor VLC with improved light uniformity.

In 2014, Wang and Su [49] used a simulated-annealing algorithm to optimize circular as well as rectangular arrays to obtain uniform illumination on the receiving surface. They also constructed an evaluation function to analyze the parameters affecting uniformity. Finally, it was concluded that the optimal radius of the array was independent of the number of light sources and produced a uniform distribution of light sources on the target surface.

Table 39.2 Research progress on the optimization of light-source positions

Year	Researchers	Study content	Progress
2009	Ca o	Proposed a quasi-sinusoidal particle-swarm algorithm to optimize the placement of light sources	Avoided the problem of local extremes, but reduced the computational accuracy of the algorithm, while keeping the number of particles constant
2015	Fu	Optimized the light-source position using a genetic algorithm	The lighting requirements were met but the placement of the light sources was too scattered for practical use
2016	Gao	Optimized the light-source position using a hybrid simulated annealing–genetic algorithm	Achieved uniform light and communication stability, but the distribution was not aesthetically pleasing and the practicality was low
2017	Liu and W	Optimized the light-source location layout using genetic algorithms for gene density	To obtain the ideal power distribution, the researcher could calculate the light-distribution scheme
2019	Zhai	Used an artificial fish-swarm algorithm to minimize the power variation at the receiving surface as an objective function and to optimize the position of the LEDs	Optimal LED-layout patterns were obtained and further improvements were made to the LED power factor
2020	Liu	Optimized LED light-source location information using multiple-swarm genetic algorithms for light-source layout models	Significantly better than layout models based on traditional genetic algorithms
2021	Li et al.	Proposed an improved artificial bee-colony algorithm to optimize the position and layout of light sources	For irregular rooms without a general light-source layout, the requirements of a standard lighting layout could also be met
2022	Qu	Proposed an evolutionary algorithm scheme for optimizing light-source positions based on maximizing the LED layout in terms of signal-to-noise coverage	Provided a new manner of thinking about the layout of interior light sources

In 2015, Chen et al. [50] used an ant-colony algorithm to find a set of power-adjustment factors to achieve illumination requirements and power uniformity on the receiving surface, thus ensuring fairness in communication between multiple users.

In 2018, Kumawat et al. [51] attempted to minimize the received power-fluctuation values in a room using a whale-optimization algorithm to obtain an optimal LED light-source layout. In the same year, Yang [52] used a bat algorithm to optimize the power of each light source, while applying it to a special indoor environment, and achieved the expected results.

In 2019, Yang et al. [53] proposed a power-optimization method with an improved ant-colony algorithm to ensure uniform lighting, found a set of power

factors to optimize them, and combined this algorithm to establish a scheme for visible-light communication. The results showed that this method improved the computing efficiency and ensured the stability of the communication environment with better convergence.

In the same year, You et al. [54] studied the layout of optical-communication light sources based on the co-optimization of position and power. They used a firework algorithm to achieve an optimal distribution of LED light-source positions and power, so that different users could receive the same, improved, quality of signal.

In 2020, Wang et al. [55] proposed an improved artificial fish-swarm algorithm to find the optimal light-source layout to search for an optimized layout that provided better communication performance. The results showed that the uniformity of the received optical power was improved and the received optical-power fluctuations were reduced, providing good communication performance by optimizing the layout of the LED light sources while meeting indoor-lighting standards.

In 2020, Zhai et al. [56] proposed a multiple-swarm genetic-algorithm-based LED light-source layout with 20 LEDs as columns to solve the problem of uneven light and power distribution by optimizing parameters, such as the position, power, and angle of the light source. This method provided a reference for LED-layout design and contributed to indoor visible-light communication systems.

In 2020, Han [57] proposed a layout design scheme for LEDs using the fast convergence and high solution efficiency of the whale-optimization algorithm to improve the LED layout, uniformity of the room lighting, and data-transmission capability. They analyzed and compared the distribution of the light sources, received power, and bit-error rate (BER) under the whale-optimization algorithm.

In 2021, Zhou et al. [58] proposed a method for jointly optimizing lamp-source locations and parameters to address the problem of uneven light-source distributions in indoor VLC systems. This method rearranged LED arrays while maintaining a constant total power and optimized the parameters in each LED array using an improved PSO algorithm. The results showed that this method could reduce the total power of the light-source layout under indoor-lighting conditions and effectively improved the light uniformity on the receiving surface, compared with conventional light-source distributions.

In 2022, Che et al. [59] proposed a square-plus-elliptical light-source layout to construct a 6 m $\times$ 5 m $\times$ 3 m room model for the problem of uneven indoor visible illumination. They used a simulated-annealing particle-swarm algorithm to optimize the spacing of the LED light sources, system-illumination uniformity, and signal-to-noise ratio to find the optimal light-source layout. The lighting and communication performance were significantly improved, compared to the traditional layout.

In 2022, Xu et al. [60] used a genetic algorithm to optimize LED emitter-unit layout parameters using a square layout to minimize 3D normalized positioning-error metrics and provided a non-uniform rectangular structure to reduce the optimized size. The results showed that this method reduced the positioning error compared to a uniform layout and provided a dense LED layout for indoor systems.

In 2022, He et al. [61] proposed a new polygonal layout, based on a rectangular planar stereoscopic space, where the illumination and signal-to-noise ratio of polygonal and circular layouts were simulated in a 5 m $\times$ 4 m $\times$ 3 m room. The location

and power of the polygonal layout were further optimized by a particle-swarm algorithm. The illumination uniformity and signal-to-noise quality factor of the optimized polygonal layout were improved.

In 2022, Zhang et al. [62] proposed an optimal layout scheme based on a parallel fully connected convolutional neural network model. They used a particle-swarm optimization algorithm to optimize the light-source position, power, angle, and other parameters. They used fully connected and parallel neural networks to model indoor VLC channels; hence, the proposed method was highly applicable to any number of light sources and any specific indoor-space size. The proposed light-source layout scheme under this model was more suitable for actual indoor scenarios, and the received signal strength was more uniform (Table 39.3).

Table 39.3 Research progress on light-source parameter optimization

Year	Researchers	Study content	Progress
2012	Ding	Custom evolved algorithms controlled the light power of individual LED lighting devices.	Enabled fully diffuse indoor visible-light communication with improved light uniformity
2014	Wang and Su	The parameters affecting homogeneity were analyzed using a simulated-annealing algorithm.	The optimum radius of the array was independent of the number of light sources and produced a uniform distribution of light sources on the target surface.
2015	Chen	An ant-colony algorithm was used to find a set of power-conditioning factors.	Provided optical power and light intensity for indoor VLC systems, ensuring fair communication between multiple users
2018	Kumawat	Received power fluctuations in the room were minimized by means of the whale-optimization algorithm.	Obtained the optimal LED light-source layout
2019	Yang	Proposed a power-optimization method with an improved ant-colony algorithm to find an optimal set of power factors.	Improved computing efficiency while ensuring a stable communication environment with good convergence
2020	Zhai et al.	Optimized light-source position, power, and angle parameters based on multiple-swarm genetic algorithms	Solved the problem of non-uniform light and power distributions
2021	Zhou	Optimized parameters in each LED array using an improved particle-swarm optimization algorithm	Reduced the total power of the light-source layout under indoor-lighting conditions and improved the light uniformity on the receiving surface
2022	Xu et al.	Optimized LED emitter-unit layout parameters using genetic algorithms	Reduced positioning errors compared to uniform layouts and provided a reference indicator for dense indoor LED layouts
2022	Zhang	Optimized the light-source position, power, angle, and other parameters using particle-swarm optimization algorithms	Could be used with any number of light sources and any specific room size for more-uniform signal strength

Summary of Research Progress

Through the previous review and analysis of research progress, this chapter presents a general outline of light-source layouts for VLC systems. According to traditional search algorithms, the light-source layout in a room can only be adjusted to meet general lighting and communication needs. However, the lighting area is not uniformly bright, while the light signal has blind spots in some areas.

Researchers have proposed a variety of light-source distribution schemes to solve such problems. While focusing on the impact of the arrangement of the light beads within the light source on the overall system, rectangular layouts of light beads have evolved to circular and hexagonal layout combinations. The research problem has been more refined through algorithms to adjust the power of each light source, emission angle, and other parameters to further improve the uniformity and stability of the power coverage.

Considering the interaction between the sources and changes in channel characteristics caused by the movement of people in the room, neural-network intelligence algorithms were integrated into the light-source layout system, which constantly adjusted the system-layout and light-source parameters according to the perceived situation to achieve real-time adaptive changes.

Visible-Light Communication Light-Source-Layout Analysis

Light-Source-Layout Principles

Indoor visible-light communication uses light-emitting diodes, such as LED lamps, as a signal source for communication; therefore, in indoor VLC, LEDs must be reasonably laid out to meet communication requirements. In general, the layout of indoor visible-light communication should consider the following aspects:

1. Light-source location: In VLC systems, the luminaires themselves act as light sources and communication devices; therefore, they must be placed in the best locations to achieve the best lighting effect and communication quality. It is recommended that luminaires for VLC systems be installed on the ceiling or walls of a room, with higher luminaires covering a wider area, and lower luminaires providing stronger communication signals.
2. Layout uniformity: Similar to the layout of visible light in a room, the layout of the light sources in a VLC system must be uniform. It is recommended to set multiple luminaires at uniform distances within the entire room and adjust the brightness control of each luminaire to improve illumination uniformity and communication signal coverage.
3. Communication obstructions: In VLC systems, obstacles significantly affect the signal, particularly when the line of sight is blocked. Therefore, when laying out light sources, the location and number of obstacles in the room must be considered to determine the best light-source-layout solution.

4. Type of light source: VLC systems can use different types of lamps, such as LED lamps and fluorescent lamps. It should be noted that different types of luminaires have different wavelengths and power characteristics, and the advantages and disadvantages of different types of luminaires and their scopes of application must be considered when designing.
5. Lighting needs: When laying out light sources, the requirements for lighting in the room must also be considered. The lighting effects should not be sacrificed because of VLC communication. It is recommended that the advantages of VLC technology be fully exploited by combining it with VLC communication technology to ensure the lighting quality.

Light-Source Spatial Layout

The spatial layout of an LED light source directly affects its irradiation characteristics, the size of the received field of view, and the received power distribution at the receiving end of the communication system. It can be determined by studying the interaction between the light source, irradiation characteristics, and indoor boundary conditions. Following years of research and development, various layout models have been developed. The following four LED-layout methods are introduced separately and the most widely used layout models are described.

Thiessen Polygon Layout

The Thiessen polygon layout is an earlier proposed model for light-source layouts for VLC systems. It is shown in Fig. 39.3, where the red triangle represents the

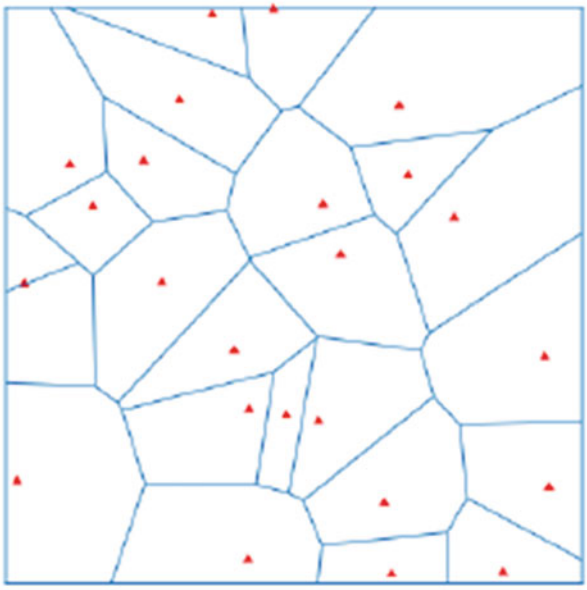


Fig. 39.3 Thiessen polygon layout [65]

location of the LED, the blue circle represents the location of the user, and the blue line represents the managed cell boundary [45].

Considering the LED positions in the room as point sets, let the number of LEDs in the room be N_l; the point set is the set of all LED positions, and the corresponding set expression is

$$I = \left\{ \left(x_1^l, y_1^l\right), \ldots, \left(x_2^l, y_2^l\right), \ldots, \left(x_N^l, y_N^l\right) \right\}. \tag{39.5}$$

Each LED serves a communication user in the corresponding area, with a total number of users N_u, and k users in each cell. The location of the user in the i^{th} cell is denoted as [63]

$$V_i = \left\{ \left\| u_i - \left(x_i^l, y_i^l\right) \right\| < \left\| u_i - \left(x_j^l - y_j^l\right) \right\|, \forall j \neq i \right\}, \tag{39.6}$$

where $\left(x_i^l, y_i^l\right)$ represents the location of the user in the i^{th} cell, $\left(x_j^l, y_j^l\right)$ represents the user location in the j^{th} cell, and V_i denotes the cell. If all cells in the room satisfy Eq. (39.6), the layout is called a Thiessen polygon layout; the distance between all users in each served area is always less than the distance between users in other cells [64]. The generation of a Thiessen polygon layout is not related to the user location [65].

Center-of-Mass Thiessen Polygon Layout

The Thiessen polygon layout was further developed, deducing the center-of-mass Thiessen polygon layout, as shown in Fig. 39.4. The geometric definition of the Thiessen polygon layout is obtained from Eq. (39.6). The user distribution is introduced to produce a center-of-mass Thiessen polygon layout, based on the user-distribution locations [66], where ρ is the location of all users, u. The center-of-mass location in the i^{th} cell is defined as

$$(x_i, y_i) = \frac{\int_{vi} u\rho(u)d\sigma}{\int_{vi} \rho(u)d\sigma}, \tag{39.7}$$

where $\left(x_i^l, y_i^l\right)$ moves the position of the LED in the i^{th} cell in the Thiessen polygon layout to the position after the center of mass. After moving all the LEDs in the cell to the position of the center of mass, it is called a center-of-mass Thiessen polygon layout. The user distribution determines the positions of the LEDs, and the positions of the LEDs in the simulation are constantly iterated according to the user distribution. The positions of the LEDs are reduced with the distance between users and finally converge.

The cost function for acquiring VLC services for a user is denoted as $F(I)$ [67]:

$$F(I) = \sum_{i=1}^{Ni} F_i = \sum_{i=1}^{Ni} \int_{vi} \left\| u_i - \left(x_i^l, y_i^l\right) \right\|^2 \rho(u)d\sigma \tag{39.8}$$

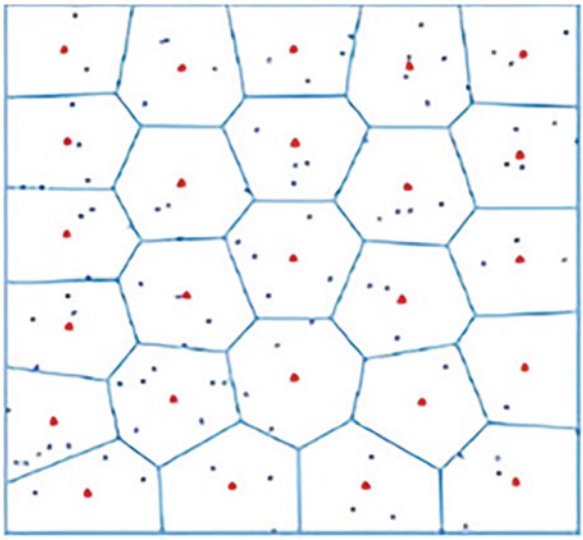


Fig. 39.4 Center-of-mass Thiessen polygon layout [45]

$F(I)$ is a function of the distance between the indoor LED and the user, and F_i represents the cost function of the horizontal distance between the LED and the user in the i^{th} cell:

$$F_i = \int_{vi} \left\| u_i - \left(x_i^l, y_i^l\right) \right\|^2 \rho(u) d\sigma \tag{39.9}$$

where $d\sigma$ denotes the integral over the area of the lighted area. In a right-angle coordinate system, $d\sigma = dxdy$, where $d\sigma$means the integration of the i^{th} cell V_i from the x and y directions. $\int_{vi} d\sigma$ and $\int_{vi} \rho(u) d\sigma$ denote the area of V_i and the number of users, respectively. A partial differentiation of the user positions in Eq. (39.9) yields

$$\frac{\partial F_i}{\partial\left(x_i^l, y_i^l\right)} = 2\int_{vi} \rho(u)\left[x_i^l, y_i^l - u_i\right] d\sigma \tag{39.10}$$

$$\frac{\partial F_i}{\partial\left(x_i^l, y_i^l\right)} = 0. \tag{39.11}$$

The user position can be expressed using Eq. (39.11), where $\left(x_i^l, y_i^l\right)$is the center-of-mass position of the Thiessen polygon layout [68]. The center-of-mass Eq. (39.11) is substituted into Eq. (39.10) and minimized; the expression is as follows:

$$F(I) = \sum_{i=1}^{Nl} F_i = \sum_{i=1}^{Nl} \int_{vi} \left\| u_i - \frac{\int_{vi} u\rho(u)d\sigma}{\int_{vi} \rho(u)d\sigma} \right\|^2 \rho(u)d\sigma \tag{39.12}$$

The above expression is the expression for the center-of-mass Thiessen polygon layout. The equal distance of the Thiessen polygon layout to discrete points in the area makes the received power distribution more uniform throughout the indoor optical communication; the irradiance can also meet the lighting requirements, and the practical application illustrates its reasonableness.

Wireless-Communication Network Layout

The wireless-communication network layout is divided into quadrangular and hexagonal layouts, as shown in Fig. 39.5; it divides the room into a square network based on the layout of the light-source locations. The quadrangular layout is widely used in indoor scenarios because of its advantages, such as uniform lighting and simple design. However, traditional layouts cannot accommodate different user distributions. In this typical LED layout, the number of light sources should be consistent.

For example, in a quadrangular layout, the number of LEDs can be 9, 16, 25, 36, 49, etc. When the user distribution changes, LEDs in other areas do not work properly because of the uneven distribution of users. This affects the communication and lighting of the entire indoor light communication system, as shown in Fig. 39.5. The hexagonal grid layout has the same advantages and disadvantages as the quadrangular grid. In the hexagonal grid layout, the number of LEDs can be 14, 23, 30, 33, 42, etc [45].

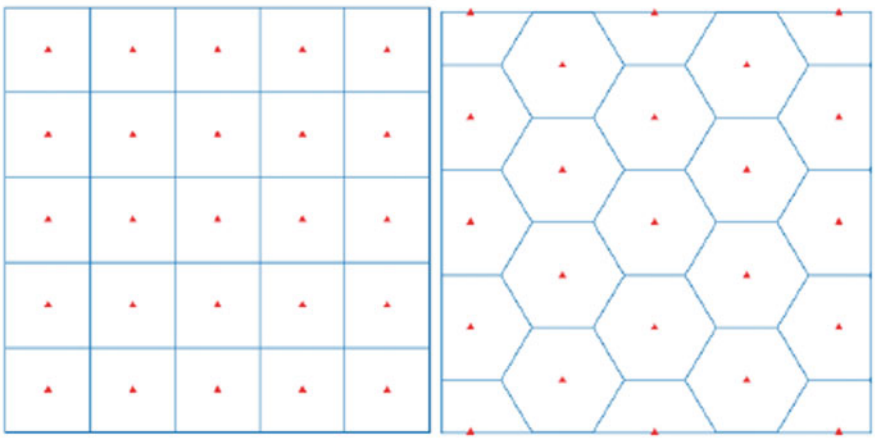


Fig. 39.5 Quadrangular and hexagonal layouts [45]

LED-Array Light-Source Layouts

Common LED-array light-source layout models include square, circular, and hexagonal [69]. Jiang et al. [32] studied the luminosity distribution of different light-source arrays composed of multiple LED beads and improved them to obtain a uniform LED-light arrangement and optimal spacing. Based on the luminosity distribution of the LEDs, a model of the luminosity distribution between individual LEDs and beads was constructed and derived for a multiarray LED [32].

Figure 39.6a–c show the square, hexagonal, and circular-array light-source layouts, respectively.

To obtain higher-quality signals, Jiang et al. [32] considered the three array-layout methods mentioned above. After continuous experimental derivation, they concluded that the relative position between the receiving surface and the LED in the visible-light communication system was fixed. Therefore, a compensation array could be added to the four directions of the annular array; after adjustment, it formed a new layout: the annular-angular-compensation array light-source layout.

Key Algorithms for Light-Source Layout Parameters

When the light-source layout is implemented under short-distance lighting conditions, it is more flexible and convenient to use algorithms to design the layout. When optimizing light-source layouts, the following algorithms have been used by scholars at home and abroad to optimize light-source layouts, which are described below.

Particle-Swarm Algorithm

The particle-swarm algorithm is a commonly used evolutionary optimization algorithm, as shown in Fig. 39.7. The particle-swarm algorithm views the candidate solutions as particles and iteratively searches for the optimal solution. Each particle adjusts its speed and position, according to its own optimal solution and that of the entire particle swarm.

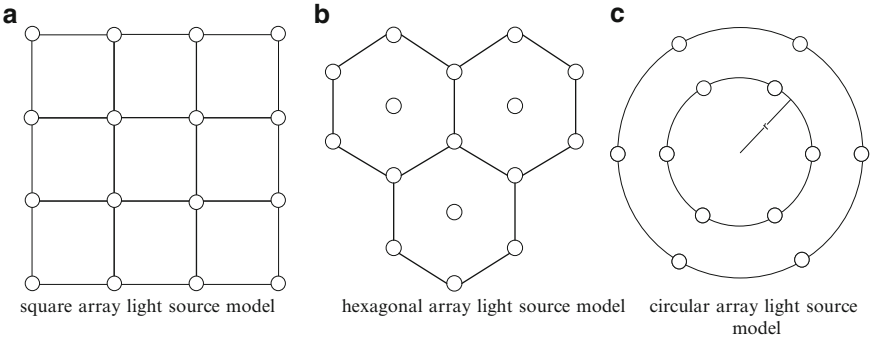


Fig. 39.6 LED-array light-source layout models [32]

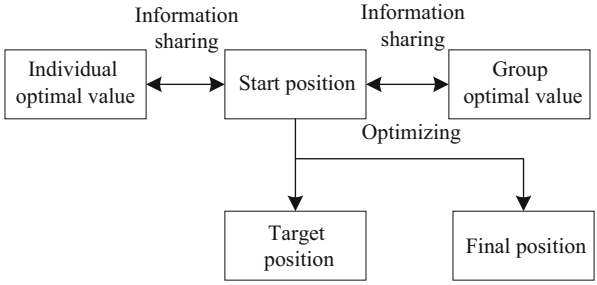


Fig. 39.7 Schematic diagram of particle-swarm optimization [22]

The particle-swarm algorithm has a good global-search capability and convergence and has been widely used in multi-objective optimization. Through continuous iteration and optimization, the particle-swarm algorithm can determine the optimal solution to a problem until the maximum number of iterations is reached. In [22], this algorithm was used to determine the optimum LED position in a rectangular plane to obtain an improved light-source layout.

An optimized layout solution based on the particle-swarm algorithm can achieve better lighting uniformity and improve the communication quality of the system, while meeting the lighting and communication requirements.

Genetic Algorithm

A flowchart of the genetic algorithm is shown in Fig. 39.8. A genetic algorithm is an optimization algorithm based on the idea of natural evolution. It searches the problem space and continuously and iteratively searches the solution space to find the optimal solution step-by-step. Genetic algorithms are widely used for optimization problems. In [38], Gao used this algorithm to optimize the light-source layout of a system using a genetic-algorithm design to obtain the LED-chip coordinates when the light distribution was optimal by optimizing the objective function.

Ant-Colony Algorithm

The ant-colony algorithm is an intelligent-optimization algorithm developed based on the foraging behavior of ants. It is highly efficient and accurate in solving various optimization problems, has good global-optimization capabilities, and can effectively handle large-scale, complex optimization problems.

Chen et al. [50] used the global-search capability of an ant-colony algorithm to find an optimal set of power-conditioning factors. The key steps of this were the selection of the objective function and state-transfer probability. To further improve the convergence of this algorithm, the intensity values were dynamically adjusted. On this basis, the light-source layout of the system was optimized, and the light uniformity in the receiving plane and the communication stability were ensured.

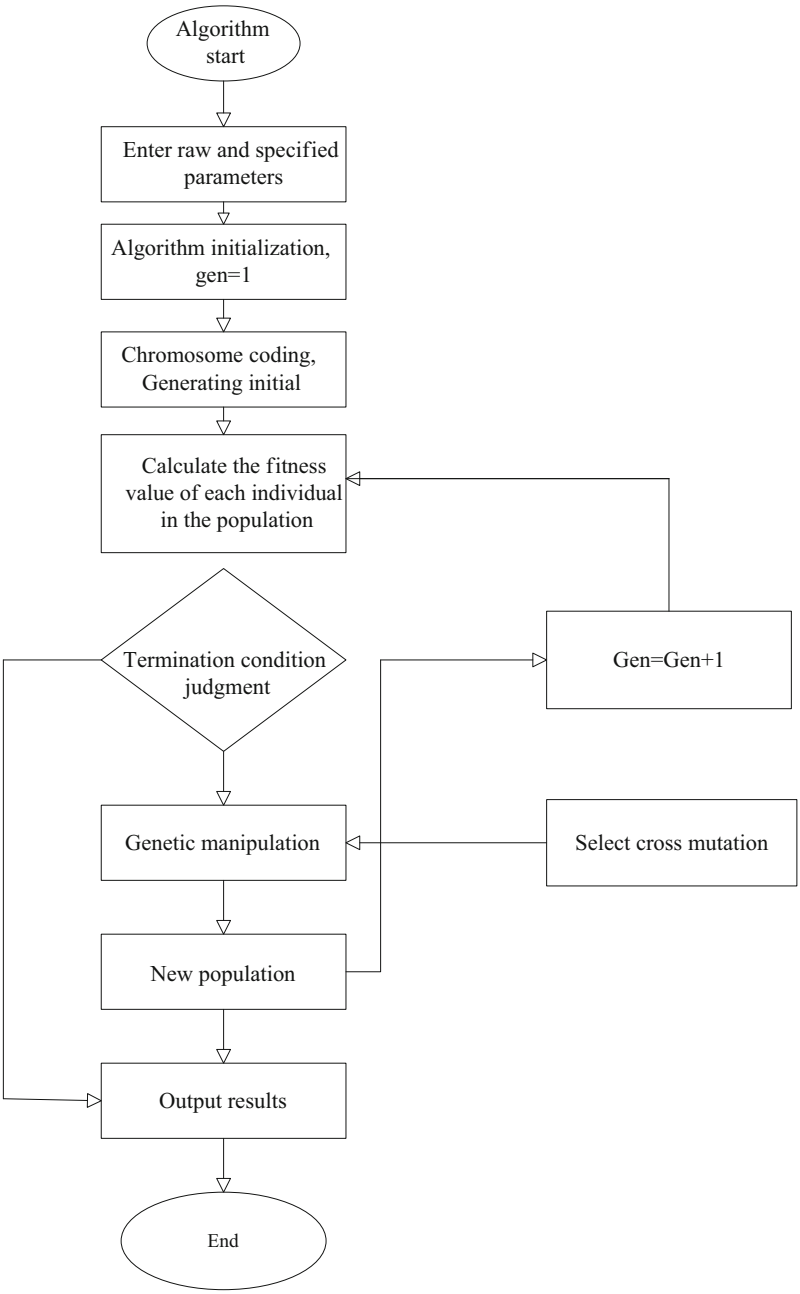


Fig. 39.8 Flowchart of a genetic algorithm [35]

Application of Intelligent Neural-Network Algorithms

In the process of optimizing the light-source layout for VLC systems, the layout must be adjusted according to the irradiation characteristics of the light-source room, received power distribution, and other parameters, and many intelligent algorithms have been introduced.

For example, the LED power factor can be adjusted in real time, based on a genetic simulated-annealing algorithm. It seeks to globally optimize the power distribution according to the cost function, ensuring the search capability and evolutionary ability of the algorithm in the process.

With the development of network technology, artificial neural-network algorithms have been applied to optimize indoor light-source layout models. Deep learning can solve the problems of long comparison and iteration processes, and the slow calculation speed of traditional algorithms.

Tan proposed the use of a feed-forward artificial neural network to optimize light-source layouts [70]. The accuracy of its prediction could be determined by comparing the optimal light-source layout scheme obtained by the genetic algorithm with that obtained by neural-network calculation. The accuracy was calculated by dividing the number of points of the same light-source arrangement by the total number of light sources.

A hybrid immunoclonal bat algorithm was proposed to optimize the layout of LED light sources in a room [71]. The optimized received optical power, signal-to-noise ratio, and illuminance distribution were compared with unoptimized and optimized data from other algorithms. Specifically, the variance in the optical power in the receiving plane was as low as 0.026 dB, after optimization using the immunoclonal bat algorithm.

Research on VLC Light-Source Layouts at Xi'an University of Technology

Received-Power Analysis

The received power is a parameter that must be considered during the VLC light-source layout; it directly determines the communication quality of the VLC system. In the design process, the received power at different locations in the room should be average and stable. Firstly, a receiving channel model was constructed to analyze the received power; the received power at a certain point in the room is shown in the figure below.

According to the reception process shown in Fig. 39.9, the received power of a point (x, y, z) is

$$P_r(x, y, z) = \sum_{j=1}^{j} H_j(0) P_t j, \tag{39.14}$$

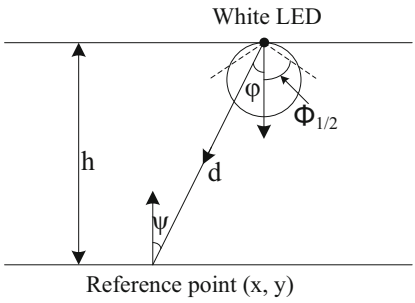


Fig. 39.9 Schematic diagram of the received power at some point in the room [72]

where P_{tj} is the total emitted power of the j^{th} array light source. $H_j(0)$ represents the channel DC gain of the j^{th} array light source at point (x, y, z), which can be expressed as [73]

$$H_j(0) = \frac{I_j(x,y,z)AT_s(\psi)\cos(\psi)}{r_j^2} rect\left(\frac{\psi}{\psi_c}\right) \tag{39.15}$$

$$rect(x) = \begin{cases} 1 & |x| \ll 1 \\ 0 & |x| > 1 \end{cases} \tag{39.16}$$

In the above equation, $I_j(x, y, z)$ is the luminous intensity distribution of the light-source array, A is the photodiode-detector receiving area, ψ is the angle of incidence, $T_s(\psi)$ is the optical-filter gain, Ψ_c is the receiver field-of-view angle, r_j is the distance between the receiver and the j^{th} array light source, and n indicates the reflection index. The optical concentrator $g(\psi)$ can be expressed as

$$g(\psi) = \frac{n^2}{\sin^2\psi_c} rect\left(\frac{\psi}{\psi_c}\right). \tag{39.17}$$

Optimal Design of a Light-Source Layout

In an indoor VLC system, each room has different dimensions and equipment, and the LED layout is arranged according to the different rooms to achieve the best communication effect. The results show that planar-array light sources have better light uniformity than linear-array light sources, suggesting a planar-array indoor VLC system.

Figure 39.10 shows a room with dimensions L × W × Z_H. For example, let there be a total of J LED-array light sources, designed as a spherical-array light source. Subsequently, the distribution of the reception rate on the plane $z = h$ is

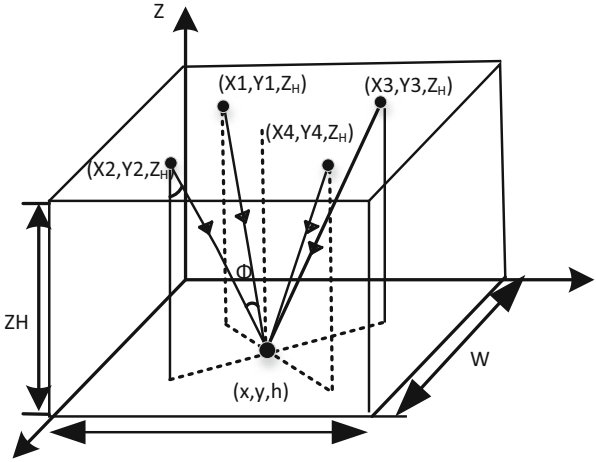


Fig. 39.10 Schematic diagram of a light-source layout [72]

analyzed. When the received-power distribution is the smallest, the LED light-source layout is the best.

Let the coordinates of the LED-array light source be (X_1, Y_1), (X_2, Y_2), ⋯, (X_J, Y_J). They can be derived from the indoor plane $z = h$ on the average value of the received power; that is, the average received power is

$$E(p_r) = \bar{p_r} = \frac{1}{s}\iint_{\Omega} p_r(x, y)dxdy, \tag{39.18}$$

where $p_r(x, y)$ is the received power value, which is distributed with equal probability in the $z = h$ plane. s is the area of plane $z = h$ in the room, and Ω represents this area.

Layout-Model Simulation Analysis

The lighting consists of an array of white LEDs [72]. To facilitate the selection of the optimal LED arrangement, the effects of the LED arrangement and received power were numerically simulated. The simulation parameters are listed in Table 39.4.

Optimal Layout Design under Different Numbers of White LED-Array Light Sources

The indoor environment is 5.0 m × 5.0 m × 5.0 m, h is 0.85 m, $X_1^* = Y_1^* = 1$, a spherical-surface array light source is used, and the design parameters are two layers, seven LEDs, $d = 0.1$ cm, and $\theta = 10^0$. The simulation results are shown in Fig. 39.11a–d.

From Fig. 39.11a–d, it can be seen that the range of received power increases from small to large as X_1 and Y_1 change from 0 m to 2 m. When $X_1 = Y_1 = 1$ m, the distribution of light power at the receiving surface is the most uniform, at which time

Table 39.4 Simulation parameters [72]

Parameter	Numerical value
Emission light power	20 [mW]
Reflective index	1.5
Transmission-power half-angle	60 [deg]
Receiver FOV	60 [deg]
PD area	1.0 [cm^2]
Optical-filter gain	1.0
Photoelectric-conversion efficiency	0.53 [A/W]
Power-transmission half-angle	60 [deg]
$W \times L \times Z_h$	$5.0 \times 5.0 \times 5.0$ [m]
h	0.85 [m]

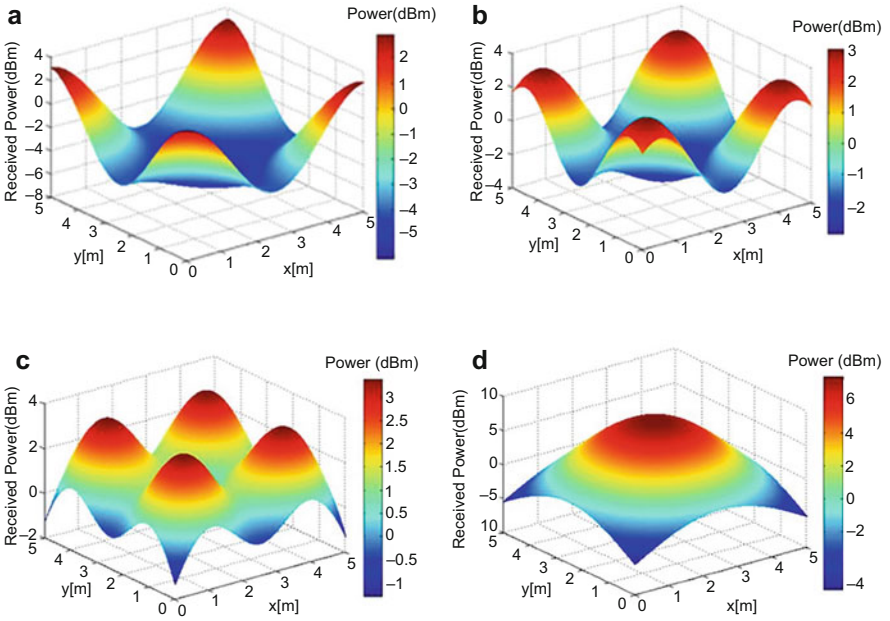


Fig. 39.11 Received optical-power distribution. (**a**) Received power distribution at $X_1 = Y_1 = 0$ m [17]. (**b**) Received power distribution at $X_1 = Y_1 = 0.5$ m [17] (min = −6.0 dBm, Max = 2.98 dBm) min = −2.9 dBm, Max = 3.0 dBm). (**c**) Received power distribution at $X_1 = Y_1 = 1$ m [17]. (**d**) Received power distribution at $X_1 = Y_1 = 2$ m [17] (min = −1.3 dBm, max = 3.4 dBm) (min = −5.6 dBm, max = 3.4 dBm) [17]

the maximum value is 3.4 dBm and the minimum value is −1.3 dBm, thus verifying the accuracy of the theoretical model.

Influence of White LED-Array Light-Source Design Parameters on the Light-Intensity Distribution

When the light-source layout is optimal, the design parameters of the LED-array light source will affect the light-power distribution in the room. The selection of key parameters, such as array-element distribution, array-element optical-axis pole

angle, and number of elements in a multi-array light source, can effectively regulate the distribution of light power and improve the light-intensity fluctuation. The following is an analysis of the impact of the design parameters of a multiarray light source on the light-intensity distribution in a room.

Figure 39.12 shows the indoor received-power distribution of the spherical-array light source. To simplify the analysis, one array light source was used. The optimal position of the array light source was in the center of the ceiling, and the number of LEDs in each layer was one, four, eight, and 16. The 3D spherical-array light sources had a height of 0.13 m; when the height of the light source is 0, the 3-D array light source becomes a 2-D array light source.

From Fig. 39.13, it can be concluded that the LED optical-axis pole angle has a great influence on the optical-power distribution: the smaller the pole angle, the greater the fluctuation and value of the optical power. Therefore, as large an optical-axis pole angle as possible should be selected while ensuring the minimum light power. Compared to 2-D array light sources, the smaller the number of layers in 3-D-array light sources, the smaller the fluctuation of the optical power distribution. The higher the number of layers of the 3-D array light sources, the larger the fluctuation of the optical-power distribution [72].

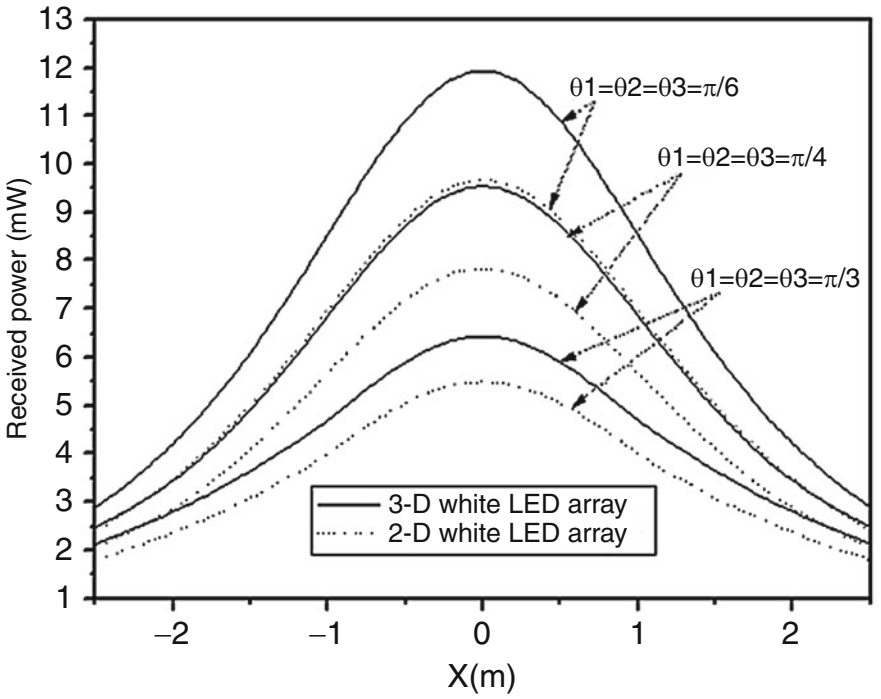


Fig. 39.12 Optical power distribution of an indoor zonal 2-m plane receiving from a spherical-array light source [72]

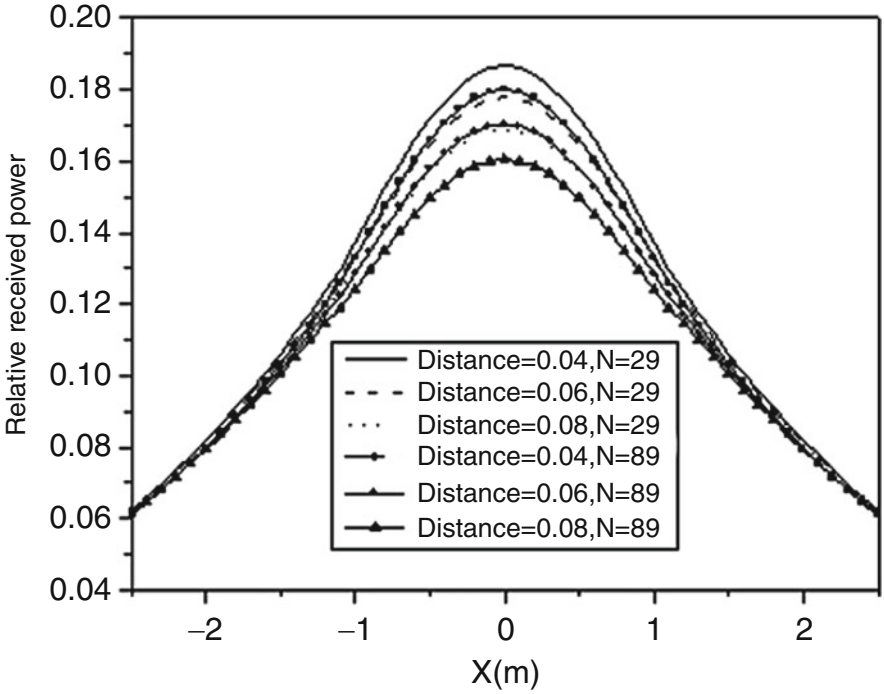


Fig. 39.13 Effect of layer spacing and number of LED elements on indoor light-intensity distribution [72]

Figure 39.13 shows the light-intensity distribution relative to the room for different LED-array unit-layer spacings and array-element numbers. Assuming that multiple white LED arrays emit an equal total power for unit 1, the luminous intensity of each array element in the 3-D spherical-surface array light-source design was spaced equally. The larger the layer spacing, the smaller the fluctuations in the light-intensity distribution. However, the layer spacing is not infinite and must be selected reasonably [72].

Summary

Visible-light communication (VLC) is an emerging communication technology whose basic principle is to use LED lights to emit data streams and transmit information through strobes that are imperceptible to the human eye. This chapter briefly introduced indoor visible-light communication systems and provided an overview of the current research status of the light-source layouts in China and abroad. It summarized and compared the existing layout models with missing light-source space and LED light beads, summarized and explained the algorithms in the

layout design process, and finally, provides an outlook on the future development of VLC and light-source layouts.

At present, the application of indoor visible-light communication still has many problems. The main light source for visible-light communication is an LED lamp. Parameters such as its brightness and the irradiation angle of divergence affect the indoor lighting situation, while the strobe characteristics also affect the quality of optical communication, human-eye health, and mental state.

Optical receiver designs, anti-interference algorithms, stable reception technology, multiple-input multiple-output technology, and more sensitive receivers will improve the quality of communication in VLC systems and are issues that need to be addressed in the process of integrating smart-living and intelligent-transport optical networks.

Currently, VLC applications are mainly focused on indoor, in-vehicle, and mining scenarios, in which the impact of the light-source distribution on the channel characteristics cannot be ignored. VLC-system light-source-layout research has produced many results, and the light-source layouts may take the following research directions in the future.

1. Multipoint access technology: Future indoor VLC systems will use multipoint access technology to support multiple users communicating simultaneously. It will use multiple light sources to provide wider coverage, higher reliability, and more fault-tolerant services. The rate and coverage area of communication systems can be further increased. This requires setting up additional light sources in the room and assigning appropriate light sources to corresponding users for communication.
2. Intelligent light-source layouts: With the development of artificial intelligence, systems can automatically optimize the position and number of light sources for intelligent management and control using machine learning and other technologies. This enables more efficient and larger data transmissions and communication.
3. Adjustable light-source layouts: In indoor visible-light communications, the brightness and color of the light source can be adaptively adjusted according to environmental changes, thereby achieving more stable and efficient communication. The layouts of future light sources will also consider these factors and more flexible, intelligent, and adjustable light-source layouts will be designed.
4. Multi-functional integrated technology: In the future, indoor VLC systems based on LED light sources will be further developed with more innovative designs and intelligent features. For example, functions such as audio, video, and data processing can be integrated into a light source for multimedia and data communication within the same light band, using spectral-analysis and demodulation algorithms. This will provide users with more convenient and richer communication services.

Overall, the layout of light sources in future indoor visible-light communications will become more intelligent, personalized, and flexible to better meet people's needs for high-speed, reliable, and secure communications.

References

1. K. Warmerdam, A. Pandharipande, D. Caicedo. Visible light communications for sensing and lighting control [J]. IEEE Sensors Journal, 2016, 13(1): 1–6.
2. Li Z, Wang ZG, Liang JY, et al. Research progress of LED visible light communication [J]. Journal of Lighting Engineering, 2023, (1): 29–40.
3. Liang JYY, Xu Yaxin, Han Mimi, et al. Research progress of heterogeneous fusion system for indoor visible optical communication [J]. Optical Communication Technology, 2023, (1): 46–50.
4. Zhang Huiping. Indoor visible light communication light source layout and its optimization [J]. Optical Communication Technology, 2017, 41(3): 20–22.
5. Ke Xizheng, Liang Jinyuan, Xu Dongsheng, et al. Research progress of pulse position modulation technology for Optical Wireless Communication class [J]. Optical Engineering, 2022, (3): 3–21.
6. Ke Xizheng, Qin Huanhuan, Yang Shangjun, et al. Nighttime background light noise model for visible light communication systems for vehicular networks [J]. Journal of Radio Wave Science, 2021, (6): 986–990.
7. Li Mengnan, Lu Wenshuai. Study on the influence of light source coherent parameters on received light intensity in spatial optical communication [J]. Journal of Chinese Academy of Electronic Sciences, 2019, 14 (5): 508–513.
8. Ruan Wei, Du Zhihua, Chen Puwang, et al. Research on optical angle selective merger diversity receiving system for visible light communication [J]. Optical Communication Technology, 2020, 44 (1): 44–48.
9. Wang J A, Che Y, Lu C, et al. LED layout for indoor visible light communication based on energy optimization [J]. Acta Optica Sinica, 2017, 37(8): 4–9.
10. Wang Xuan, Tao Tao. Modulation bandwidth study of visible light communication chips based on blue-green micro-LEDs [J]. Semiconductor Optoelectronics, 2021, 42(4): 469–473.
11. Zohaib A, Ahfay M H, Mather P J, et al. Improved BER for offset pulse position modulation using priority decoding over VLC system [J]. IEEE Wireless Days, 2019, 27(2): 27–40
12. Ren K. Research on the equalization technology of LED indoor visible light communication [D]. Lanzhou, Gansu: Lanzhou Jiaotong University, 2022.
13. Zhang Yue. Research on the integration of indoor visible light communication and positioning [D]. Xuzhou, Jiangsu: China University of Mining and Technology, 2021.
14. Ke Xizheng, Qin Huanhuan, Yang Shangjun, et al. Nighttime background light noise model for visible light communication systems for vehicular networks [J]. Journal of Radio Wave Science, 2021, 36(6): 986–990.
15. T. Komine, M. Nakagawa. Fundamental analysis for visible-light communication system using LED lights [J]. IEEE Transactions on Consumer Electronics, 2004, 50(1): 100–107.
16. Sivabalan A, John J. Improved power distribution in diffuse Indoor Optical Wireless systems employing multiple transmitter configurations [J]. Optical & Quantum Electronics, 2006, 38(8): 711–725.
17. Ding DQ, Ke Xizheng, Li JX. Light source layout design and simulation study of VLC system [J]. Optoelectronic Engineering, 2007, (1): 131–146.
18. Nguyen H Q, Choi J H, Kang M, et al. A MATLAB-based simulation program for indoor visible light communication system [C]//7th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2010), Newcastle, United Kingdom: IEEE, 2010: 537–541.
19. Taparugssanagorn A, Siwamogsatham S, Pomalaza-Ráez C. A miso uca beamforming dimmable led system for indoor positioning [J]. Sensors, 2014, 14(2): 2362–2378.
20. Chen Ran. Mathematical model of indoor lighting communication system and its simulation [D]. Nanjing, Jiangsu: Nanjing University of Posts and Telecommunications. 2013.
21. Wang L, Guo Mao-Tian, Tian Hui. A visible light communication light source layout model and performance analysis [J]. Laser Journal, 2016, 37(3): 92–94.

22. Chen, S., Ling, L.I., Huang, J.W., et al. Indoor visible light communication LED light source layout and optimization [J]. Optical Communication Technology, 2018, 42(10): 56–59.
23. Zhao L, Peng K. Optimization of light source layout in indoor visible light communication based on white light-emitting diode [J]. Acta Optica Sinica, 2017, 37(7): 13–20.
24. Chen C, Zhong W D, Wu D. On the coverage of multiple-input multiple-output visible light communications [Invited] [J]. Journal of Optical Communications and Networking, 2017, 9(9): D31-D41.
25. Garg P, Gupta A. Comparative analysis of hexagonal VLC nodes deployment schemes [C]//4th International Conference on Signal Processing, Computing and Control(ISPCC), Solan, India. IEEE, 2017: 368–372.
26. Zhao L, Peng K. Optimization of indoor visible light communication light source layout based on white light-emitting diodes [J]. Journal of Optics, 2017, 37(7): 49–57.
27. Wang Jiaan, Che Ying, Lv Chao, et al. Energy-optimal LED layout for indoor visible light communication based on energy optimization [J]. Journal of Optics, 2017, 37(8): 4–10.
28. Zhao Zixu, Song Xiaoqing, Jia Shengjie, et al. Optimization of light source layout for visible light communication system in special vehicles [J]. Infrared and Laser Engineering, 2017, 20(01): 230–250.
29. Zhao L, Zhu T, Liu ZG, et al. A toroidal light source layout model that balances illumination and communication [J]. Optoelectronic Engineering, 2018, 45(7): 4–9.
30. R. Sharma, A. C. Kumari, M. Aggarwal. Down-link performance of an LED based visible light communication system through dynamic controlof semi-angle [J]. Physical Communication, 2019, 33(2): 222–237.
31. Chen, Quanrun, Zhang, Tao. Optimization of light source layout and performance analysis of indoor visible light communication system [J]. Journal of Optics, 2019, 39(4): 87–98.
32. Jiang XM, Dai SHA, Wu Hao, et al. Research on the optimal layout of light sources for point-to-point visible light communication systems [J]. Communication Technology, 2020, 53(3): 543–550.
33. Wang Jia'an, Lv Chao, Gu Yixingwen, et al. Research and performance analysis of LED layout for indoor visible light communication [J]. Information Communication, 2021, 34(1): 21–24.
34. Cao S, An C Y, Zhang Y. An improved particle swarm algorithm and its application to optimal layout of light sources [J]. Microelectronics and Computers, 2009, 26(7): 69–72.
35. Guan R, Wang J Y, Wen Y P, et al. PSO-based LED deployment optimization for visible light communications[C]//2013 International Conference on Wireless Communications and Signal Processing. Hangzhou, China: IEEE, 2013: 1–6.
36. Fu YQ. Optimization of visible light communication LED array distribution using genetic algorithm [J]. Journal of Huaqiao University (Natural Science Edition), 2015, 36(6): 632–648.
37. Krommenacker N, Vasquez O C, Alfaro M D, et al. A self-adaptive Cell-ID positioning system based on visible light communicationsin underground mines[C]//2016 IEEE International Conference on Automatica, IEEE, 2016: 1–7.
38. Gao Mingguang. Research on optimizing light source layout based on LED indoor visible light communication system [D]. Beijing: Beijing University of Technology, 2016.
39. Ajmani M, Sinanović S, Boutaleb T. Optimal beam radius for LED-based indoor positioning algorithm[C]//2016 International Conference for Students on Applied Engineering (ICSAE). IEEE, 2016.
40. Liu H, Wang X, Chen Y, et al. Optimization lighting layout based on genedensity improved genetic algorithm for indoor visible light communications [J]. Optics Communications, 2017, 39(02): 76–81.
41. Liu Yun. Research on light source layout and positioning application of LED indoor visible light communication [D]. Guangzhou: South China University of Technology, 2018.
42. Zhai Changxin, Zheng Ru, Liu Chang, et al. Optimal design of light source for indoor visible light communication system [J]. Optical Communication Technology, 2019, 43(8): 11–15.
43. Liu H, Zhai C.X., Wen Y.Y., et al. Optimal layout model of light sources for visible optical communication systems [J]. Optical Engineering, 2020, 47(07): 63–70.

44. Ma Yiming, Duan Mingming, Hao Xiangyin, et al. Layout study of visible light wireless communication system in civil aviation cockpit [J]. Science and technology wind, 2020, (06): 128.
45. Li Jiangtao. Study on the optimization of indoor visible LED layout [D]. Zhenjiang, Jiangsu: Jiangsu University, 2021.
46. Li Bo, et al. Light Source Layout Optimization Strategy Based on Improved Artificial Bee Colony Algorithm [J]. Mathematical Problems in Engineering, 2021, 20(3): 16–30
47. Qu Shupeng, Liu Defeated. LED layout optimization scheme [J]. Optical Communication Technology, 2022, 46(2): 45–50.
48. Ding J.P, Ji Y.F. Evolutionary algorithm-based optimisation of the signal-to-noise ratio for indoor visible-light communication utilising white light-emitting diode [J]. IET Optoelectronics, 2012, 6(6): 307–317.
49. Wang Jiawen, Su Zuoping, Yuan Zhijun, et al. Illumination uniformity in LED array modularity [J]. Journal of Photonics, 2014, 43(8): 22–28.
50. Chen Y,Yin F,Liu Huanlin,et al. Optimization of optical power in LED communication using ant colony algorithm [J]. Optoelectronics-Laser, 2015, 26(3): 462–467.
51. Kumawat I R, Nanda S J, Maddila R K. Positioning LED panel for uniform illuminance in indoor VLC system using whale optimization [M]. Springer, Singapore: Optical and Wireless Technologies, 2018: 131–139
52. Yang M. Research on the optimization of power distribution based on bat algorithm in visible optical communication [D]. Changchun, Jilin: Jilin University, 2018.
53. Yang Zhiqun, Liang Jiran, Mao Luhong. Study of ant colony algorithm for optimizing power uniformity of indoor visible light communication [J]. Advances in Lasers and Optoelectronics, 2019, 56(22): 41–47.
54. You Chunxia, Hu Qingsong, Li Shidang, et al. Visible light communication light source distribution at coal mine working face based on synergistic optimization of position and power [J]. China Laser, 2019, 46(04): 265–272.
55. Wang J A, Xu A C, Ju J T, et al. Optimization lighting layout of indoor visible light communication system based on improved artificial fish swarmalgorithm [J]. Journal of Optics, 2020, 22(3): 35–49.
56. Zhai C X, Liu H, Wen Y Y, et al. Multipopulation genetic algorithm-optimized LED layout in a visible light communication system [J]. Optical Engineering, 2020, 59(6): 18–46.
57. Han Yu, Yue Dianwu. Indoor visible light communication LED layout based on whale optimization algorithm [J]. Optoelectronics, 2020, 40(3): 201–208.
58. Zhou X, Zhang YR. Lamp source layout for indoor visible light communication system based on improved particle swarm algorithm [J]. Optical Communication Technology, 2021, 45(1): 58–62.
59. Liu C, Tang Y, Yan W, et al. Optimizing the Light Source Layout of the Indoor Visible Light Communication System [J]. IEEE Access, 2022, 10(7): 27223–27235.
60. J. Xu, N. Huang and C. Gong, Indoor 3-Dimensional Visible Light Positioning Based on Smartphone Camera: Error Metric and LED Layout Optimization [J]. IEEE Photonics Journal, 2022, 14(6): 1–13.
61. Pei H, Jing L, Tong Z, et al. Layout and optimization of LED light source for indoor visible light communication [J]. Microwave and Optical Technology Letters, 2023, 65(2): 710–716.
62. Zhang Huiying, Lu Yuxi, Liang Yu, et al. Optimal lamp source layout for indoor visible channels based on parallel fully connected convolutional neural network models [J]. Journal of Optics, 2022, 42(13): 43–55.
63. Q. Du, V. Faber, and M. Gunzburger. Centroidal Voronoi tessellations: applications and algorithms [J]. SIAM Review, 1999, 41(4): 637–676.
64. M. Iri, K. Murota, and T. Ohya. A fast Voronoi-diagram algorithm with applications to geographical optimization problems [J]. Computational Statistics & Data Analysis, 1984, 62(5): 273–288.
65. J. C. Hateley, H. Wei, and L. Chen. Fast methods for computing centroidal voronoi tessellations [J]. Journal of Scientific Computing, 2015, 63(1): 185–212.

66. Y. Liu, W. Wang, B. Levy. On centroidal Voronoi tessellation—energy smoothness and fast computation [J]. ACM Transactions on Graphics (To G), 2009, 28(4): 1–7.
67. E. Abaya and F. Wise. Convergence of vector quantizers with applications to optimal quantization [J]. SIAM Journal on Applied Mathematics, 1984, 44(8): 183–189.
68. G. Celeux and G. Govaert. A classification EM algorithm for clustering and two stochastic versions [J]. Computational Statistics & Data Analysis, 1992, 14(8): 315–339.
69. Ivan M, Maximino A, Tzonchev R I Designing Light-emitting Diode Arrays for Uniform Near-field Irradiance [J]. Appl. Opt., 2006, 45(10): 2265–2272.
70. Tan Mengyuan. Dynamic optimization of light source layout for indoor visible light communication system based on deep learning algorithm [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2020.
71. Che, Hui. Optimization of light source and safety performance of indoor visible light communication system [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2022.
72. Ding Deqiang. Research on indoor visible light communication technology [D]. Xi'an: Xi'an University of Technology, 2011.
73. Liu Jianbin, Li Peixian, Hao Yue. Research and progress of high brightness GaN-based blue and white LEDs [J]. Journal of Quantum Electronics, 2005, 22(5): 673–679.

Resource-Allocation Strategy for VLC/RF Systems

40

Contents

Introduction 1496
Domestic and Foreign Research Progress 1497
System Model 1501
  Downlink Communication 1502
  Uplink Communication 1506
  OFDMA System Model 1506
  Queue Model 1508
Optimization Theory 1509
  AP Allocation 1511
  VLC AP Joint Subcarrier and Power Allocation 1512
  RF AP Joint Resource Block and Power Allocation 1515
Algorithm Simulation Results Analysis 1517
Conclusion 1523
References 1523

Abstract

To address the problem of resource-allocation strategies that limit the system transmission rate in visible light communication and radio frequency (VLC/RF) systems, a new resource-allocation method is proposed and orthogonal frequency-division multiple-access (OFDMA) technology is introduced. Thus, the capacity of the communication system is effectively increased. In the VLC/RF system model based on OFDMA, the Lyapunov optimization method is used to transform the time-averaging problem into a series of single-slot online problems. To reduce its computational complexity, the optimization problem is decomposed into three independent subproblems. The Lagrange relaxation method and convex-optimization theory are used to solve the subproblems to maximize the average system transmission rate. The Lyapunov drift is used to ensure the stability of the system. The simulation verified that the Lyapunov optimization algorithm does not require iterations, which greatly improves the optimization speed. In addition, the simulation results show that using OFDMA technology in

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_40

the indoor VLC/RF system effectively improves the system capacity, while solving the optimization problem. The proposed resource-allocation strategy effectively balances the system queue length and transmission rate, and maximizes the average system transmission rate while ensuring system stability. The Lyapunov optimization algorithm is comparable to other algorithms in terms of system stability and transmission rate, thereby proving its effectiveness.

Keywords

Indoor visible light communication · Orthogonal frequency-division multiple access · Resource-allocation strategy · Lyapunov optimization algorithm · Average transfer rate

Introduction

Future 6G networks will support data rates of up to 1 Tbit/s, while also facing the challenge of coping with a large number of mobile devices [1]. However, it is difficult for limited radio spectrum resources to meet the needs of massive amounts of data transmissions transmissions. Visible light communication (VLC) combines the advantages of lighting and communication with hundreds of high-frequency bandwidth spectra that do not require certification [2], alleviating the pressure on the insufficient radio frequency (RF) spectrum.

VLC has several advantages, such as a high data rate, low power consumption, high security, and low latency [3]. Therefore, VLC is considered a promising complementary technology for emerging wireless communication. Data transmissions in VLC systems are achieved through intensity modulation on the transmitter side and direct detection on the receiver side (IM/DD).

As with other high-frequency communication technologies, signal propagation at this frequency is highly susceptible to obstacles. Therefore, data transmissions between the transmitter and receiver in VLC require a line of sight (LOS) path. When there is a non-line of sight (NLOS) transmission between the VLC access point (AP) and the user AP, the VLC performance degrades sharply. The coverage area of the VLC AP is limited, therefore transmission reliability has become the main concern for VLC systems.

Conversely, in the case of NLOS transmissions, RF systems can provide a wider coverage area [4]. Therefore, VLC/RF heterogeneous networks have received considerable attention in recent years and have significantly improved in terms of the sum rate [5], spectral efficiency [6], power consumption [7], and energy efficiency [8]. However, the introduction of VLC into RF systems faces problems of access-point and resource allocation.

For VLC/RF heterogeneous systems, access-point and resource allocation are two important issues to be resolved before the user transmits data and are also important factors affecting the system transmission rate. At present, most research on VLC/RF

hybrid systems is focused on power allocation, and few related studies have been conducted on how to reasonably allocate APs to users.

Ma et al. [9] developed a two-stage access-point selection (APS) method based on a fuzzy logic algorithm to maximize the energy efficiency under minimum transmission-rate requirements and power constraints. Obeed et al. [10] proposed an iterative algorithm to assign APs to users, and derived a new efficient algorithm to find the optimal dual variable through dual-variable correlation. It allocates the power of each AP to connected users, to maximize the total system capacity and improve the system fairness. Pratama et al. [11] proposed a bandwidth aggregation protocol to optimize the system throughput. However, while ensuring system stability, challenges still exist in effectively and dynamically allocating AP, bandwidth, and transmission power, and jointly optimizing the system performance. First, different types of APs can make system design difficult. In general, binary allocation variables and power allocation cause the resource-allocation problem to become a mixed-integer nonlinear-programming (MINP) problem, which is a nonconvex problem with high computational complexity. Second, resource allocation should be carried out under random channel conditions. As a joint optimization problem, it is usually necessary to optimize multiple parameters simultaneously. Thus, the resource-allocation optimization problem is affected by many factors and is difficult to solve.

Domestic and Foreign Research Progress

Visible light communication is a key research direction in optical wireless communication (OWC). An OWC system relies on light radiation to transmit information in free space with wavelengths ranging from infrared (IR) to ultraviolet (UV), including the visible spectrum. Infrared with wavelengths ranging from 780–950 nm is the best choice for IR indoor optical wireless systems. Within this range, it is easy to obtain a low-cost light source, and the frequency band is consistent with the peak sensitivity of inexpensive photodiodes.

In 1996, the IR system proposed by March and Khan [12, 13] achieved a data rate of 50 Mbit/s. In 2000, Carruther and Kahn [13] developed an angular-diversity infrared communication system that could achieve a data rate of 70 Mbit/s and proved that the system could theoretically achieve a communication rate of 100 Mbit/s.

Subsequently, visible light communication technology using white light-emitting diodes (LEDs) was proposed and has received increasing attention from academia and industry. As early as 1999, Pang et al. [14] proposed to modulate information onto an LED. Owing to the high switching-cycle frequency of the LED, the human eye cannot distinguish it, making the LED appear to be constantly illuminated without affecting its lighting function. Subsequently, an experimental system was developed by Kwan to use LEDs for synchronous signals and communication. Line of sight (LOS) and IM/DD were used in LED traffic lights. The transmission

distance of the experimental system reached 20 m [15], which was the beginning of VLC research.

Subsequently, many researchers conducted extensive research on VLC and explored the influence of different factors on VLC systems. In 2000, Tanaka et al. [16] introduced a wireless home link (WHL) into visible light communication and proposed a wireless VLC system based on white LEDs. The communication system was simulated using subcarrier modulation and binary phase-shift keying (BPSK). The influences of the optical-link parameters and field of view (FOV) on the communication system were discussed. Conclusions were drawn through the simulation experiments.

When only an LOS link is used, the noise bandwidth determines the communication performance of the system, and intersymbol interference (ISI) has a greater impact when considering multipath dispersion. The larger the FOV, the more paths detected by the receiver, which leads to a decrease in performance; therefore, it is necessary to select the appropriate FOV.

The following year, Tanaka et al. studied a three-color LED and conducted simulation experiments, which increased the VLC data-transmission rate to 400 Mbit/s [17]. In 2002, Fan and Komine et al. [18] studied the effects of reflective optical communication in VLC communication systems. The study found that, owing to the multipath effect, the communication interference near the wall increased and the communication performance characteristics worsened. Therefore, the system transmission rate could be improved by reducing the interference of multipath optical signals, as well as using parallel transmissions.

Many organizations and institutions have provided VLC standards and practical models, which have made significant contributions to its rapid development. In 2003, several scientific research institutions and enterprises in Japan established the Visible Light Communication Association [19]. It attracted the interest of a large number of research institutions and enterprises, and greatly promoted the commercial development of visible light communication. In addition, it has studied other uses of VLC, such as applications in indoor positioning, and promoted VLC development. In 2004, the Visible Light Communication Association announced the practical use of white LEDs in communication, and released an indoor VLC communication system and its basic principles and channel models to the public [20].

In 2006, Hidemitsu et al. [21] proposed a new modulation method that used reverse pulse-position modulation (RPPM) and subcarrier RPPM for VLC, which improved the brightness of LEDs and weakened the influence of background light on communication. In the same year, Afgani et al. [22] verified the feasibility of orthogonal frequency-division multiplexing (OFDM) technology in visible light communication. When the transmitter and receiver were 1 m apart and the bit-error rate (BER) was 10^{-3}, the communication performance of a single LED was improved. Moreover, it was verified that reducing the bit-error rate enabled a longer communication distance.

In 2008, Hashemi et al. [23] used OFDM technology in a VLC system to reduce the intersymbol interference (ISI). Because ISI is caused by the multipath effect, the system improved the LOS and diffused multipath links. Binary phase-shift keying

(BPSK), quadrature phase-shift keying (QPSK), and quadrature-amplitude-modulation (QAM) schemes were applied to OFDM, and their communication performance was analyzed. The simulations achieved a transmission rate of 100 Mbit/s.

In 2009, Langer et al. [24] presented a practical system based on indoor VLC using on–off keying (OOK) to achieve a transmission rate of 125 Mbit/s over a distance of 5 m. They presented a visible light link using discrete multitone (DMT) to realize digital signal processing, achieving a transmission rate of more than 200 Mbit/s.

In 2009, Vucic et al. [25] demonstrated a VLC system using DMT signal modulation. The transmission rate exceeded 200 Mbit/s at the position where the receiver could obtain approximately 1100-lx illumination. The transmission rate reached 180 Mbit/s at the position where the receiver could obtain approximately 550-lx illumination, and the transmission rate reached 100 Mbit/s at the position where the receiver obtained approximately 180-lx illumination.

In 2010, Vucic et al. proposed a wireless point-to-point communication link based on white LEDs and avalanche photodiodes. By effectively utilizing the modulation bandwidth and high signal-to-noise ratio (SNR) of the LED, and using QAM and DMT technologies, the data rate was increased to 513 Mbit/s [26]. In 2012, Vucic et al. proposed a VLC system based on wavelength-division multiplexing (WDM) and DMT modulation. In this system, an RBG-type LED was used to achieve a communication rate of 830 Mbit/s [27].

In addition to studying modulation methods, many researchers have focused on device performance and combined multiple modulation techniques to improve the performance of VLC systems. In 2010, Ahmad et al. [28] presented a high-speed VLC system that used an imaging receiver. The system used multiple-input multiple-output (MIMO) OFDM technology, a 2×1 LED emission matrix, and a 3×3 photodetector array to achieve a transmission rate of 220 Mbit/s within 1 m. However, the coverage of the system was small and could only cover an area of 8 cm $\times$ 9.7 cm.

In 2013, Zhang et al. [29] used complementary metal oxide semiconductor (CMOS)-integrated GaN-based blue micron-sized LEDs for communication. They achieved a maximum data-transmission rate of 1.5 Gbit/s in free space using 16 independent parallel data streams. In 2014, Tsonev et al. [30] presented a VLC system based on QAM-OFDM. The system used a high-modulation-bandwidth GaN LED, combined with equalization and adaptive data-loading technology, to achieve a transmission rate of 3.5 Gbit/s.

In 2016, Lu et al. [31] proposed and demonstrated a high-speed wireless VLC system with a modulation bandwidth of 1 GHz. Through various techniques, such as forward error correction, OFDM, and bit power-loading algorithms, a rate of 11.1 Gbit/s was achieved in a 1.2-m free space.

In recent years, many researchers have applied new technologies, such as multi-color modulation and nonorthogonal multiple access (NOMA), to VLC and have achieved good results. In 2019, Ye et al. [32] proposed a 4×4 color-polarization multiplexing method. The available bandwidth of blue and green LEDs for signal modulation was greater than 465 MHz. Under the extremely low illumination of

6.9–136.1 lx indoors, the VLC system provided a 1.7–2.3 Gbit/s OOK service rate in a free-space transmission length of 1–4 m. In addition, the measured BER of each LED was less than 3.8×10^{-3} and the system performance was good.

In 2020, Gunawan et al. [33] proposed a color-shift keying embedded DC optical OFDMA scheme for a VLC system and presented a demonstration. The VLC system not only provided nonblinking color-shift keying signals but also provided additional modulation sizes by including high-data-rate OFDM signals. In 2021, Lin et al. [34] studied signal-demodulation and signal-recovery methods for the NOMA-VLC system. They proposed a NOMA-VLC demodulator based on a convolutional neural network, so that the signal could be simultaneously compensated and recovered. This method effectively improved the system transmission performance, user's channel gain, and user's transmission rate, while satisfying the bit-error rate limit.

Resource allocation is also an important method for improving the transmission rate of VLC systems. Different resource-allocation algorithms have been proposed for different network environments and system-performance requirements.

Abdelhady et al. [35] proposed a low-complexity, nonconvex, joint time and power-allocation method to maximize the spectral efficiency (SE) and minimize the rate constraint of the algorithm. In a VLC system with time-division multiple access (TDMA), the successive convex approximation (SCA) and Karush–Kuhn–Tucker (KKT) methods were used to simplify the optimization problem. Subsequently, an SCA problem-solving method based on the interior-point method was proposed to reduce the complexity of the optimization process.

Considering the problems of inter-cell interference (ICI) and LOS link congestion in VLC systems, Aboagye et al. [36] proposed an optimization method based on energy efficiency (EE). Considering the minimum rate limit of users, the joint subchannel and power-allocation problem for maximizing the EE appeared. To address this problem, a low-complexity iterative solution based on the fractional-programming theory and a quadratic-transformation method was proposed.

Fu et al. [37] studied the power-allocation problem of OFDM-based non-orthogonal multiple-access (NOMA) VLC systems. Considering any number of multiplexed users, an enhanced power-allocation scheme was proposed to optimize the subcarrier power, which greatly improved the subcarrier loss rate and total data-transmission rate of the system.

In VLC/RF heterogeneous networks, resource allocation is also an important way to improve the system performance. To maximize the average system throughput, Li et al. [38] studied the load-balancing problem in VLC/RF heterogeneous networks and proposed centralized and distributed schemes. Based on user fairness, these two schemes provided the AP service-time ratio and access scheme for each user. These two schemes solved the time-proportion problem using nonlinear programming and average allocation, respectively. In the distributed scheme, the load-scheduling problem was solved using the Lagrange duality method.

Mohanad et al. [39] proposed a joint power-allocation and load-balancing algorithm to maximize the system throughput. The method allocated power and bandwidth equally among users accessing different access points. According to the established throughput utility-function expression, the Lagrange multiplier method

was used to optimize the load and power. This scheme allowed access to a large number of users and improved the system throughput and fairness.

Kafafy et al. [40] proposed a joint power and bandwidth allocation scheme for improving the power efficiency in VLC/RF heterogeneous networks. This method iteratively optimized the allocation of access points, as well as the power and bandwidth allocated to users by the APs to improve the power efficiency of the system.

In this study, we investigated the performance of VLC/RF systems based on orthogonal frequency-division multiple-access (OFDMA) technology. The system consists of multiple VLC APs and one RF AP that work together to serve multiple users. Using the OFDMA scheme, resources can be simultaneously divided into the time and frequency domains, and the divided time-frequency block is called a resource block (RB). For each resource block, the smallest indivisible block is called a resource unit (RU).

OFDMA technology allows multiple users to share all the subcarriers of a system; however, each subcarrier is limited to one user in a given time slot. If other users need to use this subcarrier, they must queue up for the next time slot, which reduces the system efficiency. Allocating system resources to different users, rather than simply allocating subcarriers or time slots, can improve the efficiency and flexibility of the entire communication system.

In addition, OFDMA can flexibly select channels with different frequencies for users and adaptively allocate resources, thereby significantly improving the spectral efficiency and transmission rate of the system. At present, in the field of VLC, OFDMA research mainly focuses on spectral efficiency; however, its combination with the system transmission rate and stability has not been thoroughly investigated.

In this study, in an indoor VLC/RF system based on OFDMA technology, Lyapunov optimization was used to transform the time-averaging problem into a single time-slot problem, which was divided into three independent subproblems to optimize a single time slot.

System Model

The VLC system model uses VLC technology to transmit signals. In this model, the downlink uses VLC and RF technologies. VLC is complementary to RF in transmitting information in the form of visible light to the receiver through LED lights. The uplink uses RF technology to transmit the user's feedback signal back to the transmitter using radio waves.

This model takes advantage of the high frequency and large bandwidth of visible light, making data transmissions more efficient and reliable. In addition, by using RF technology as an uplink, users can easily interact and provide feedback to the system.

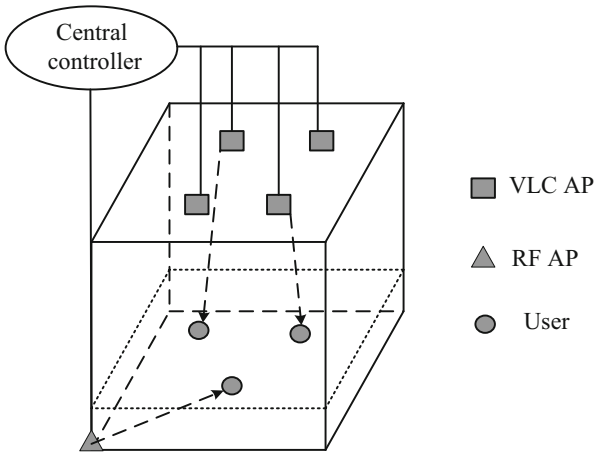


Fig. 40.1 Indoor VLC/RF heterogeneous system model

Downlink Communication

A model of the indoor VLC/RF heterogeneous system is shown in Fig. 40.1. It contains one RF AP, L VLC APs, and U users. The VLC AP was mounted on the ceiling, assuming that the RF AP could cover the entire room area. All APs were connected to a central controller responsible for collecting user feedback, scheduling, associations, and resource allocation. The users were evenly distributed on the receiving plane and equipped with VLC and RF receivers.

The set of APs is $L = L_{\mathrm{VLC}} + L_{\mathrm{RF}} = \{0, 1, 2,\ldots, L\}$, where $L_{\mathrm{VLC}} = \{1, 2,\ldots, L\}$ and $L_{\mathrm{RF}} = \{0\}$, representing the sets of VLC APs and RF APs, respectively. The user set is $U = \{1, 2,\ldots, U\}$. In this study, the communication system was divided into T time slots, and the duration of each time slot was t s. It was assumed that all user locations remained unchanged for a short period of time.

For network selection, the variable $a^{(u,\ l)}(t)$ is introduced, where $a^{(u,\ l)}(t) = 1$ indicates that user u uses AP_l for access in time slot t. Otherwise, $a^{(u,\ l)}(t) = 0$. As a single user can only be served by one AP in a time slot, the following constraints are obtained:

$$\sum_{l \in L} a^{(u,l)}(t) = 1, a^{(u,l)}(t) \in \{0, 1\}, \forall l. \tag{40.1}$$

VLC Channel Model

For indoor VLC systems, directional direct-view communication links are generally used [12]; that is, the transmitter and receiver of the VLC system are within their respective field-of-view (FOV) ranges, as shown in Fig. 40.2. For the VLC system investigated in this study, the LOS channel gain between the AP_l and user u is

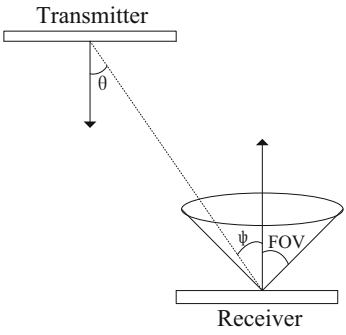


Fig. 40.2 Indoor VLC downlink LOS link model

$$h^{(u,l)} = \begin{cases} \dfrac{(m+1)A_p}{2\pi d^2} \cos^m(\theta) \cos(\varphi) T_l G_l, (\varphi \leq \text{FOV}) \\ 0, (\varphi > \text{FOV}) \end{cases}, \tag{40.2}$$

where $m = -\ln 2/\ln(\cos(\text{FOV}))$ is the Lambertian order, FOV is the half angle of the light-source transmitting power, A_p is the effective area of the receiver photodiode, d represents the distance between AP_l and user u, and θ and φ represent the irradiation and incidence angles, respectively. T_l represents the gain of the optical filter and G_l represents the gain of the concentrator.

In VLC systems, signal-transmission methods for intensity modulation and direct detection (IM/DD) have been adopted. Because the LED operates in a linear region, the output optical power of the modulated signal is proportional to the input voltage, ensuring that the signal transmitted to the receiving end is positive and steady. In addition, the average emitted electrical power P_e is satisfied by the average emitted optical power P_o:

$$P_e = \beta (P_o)^2, \tag{40.3}$$

where β is the DC bias factor of the VLC AP.

Assuming that there are K subcarriers in the VLC system, $b_k^{(u,l)}(t)$ represents the distribution coefficient of user u with the K subcarriers of VLC AP_l in time slot t, and $b_k^{(u,l)}(t) = 1$ indicates that AP_l assigns subcarrier k to the user for communication in time slot t. Otherwise, $b_k^{(u,l)}(t) = 0$.

The signal-to-noise ratio on the kth subcarrier between user u and VLC AP_l is

$$\xi_k^{(u,l)} = \frac{\left| \varepsilon \beta p_{k,o}^{(u,l)}(t) h_k^{(u,l)}(t) \right|^2}{\sigma_{\text{VLC}}^2} = \frac{p_k^{(u,l)}(t) \left| \varepsilon h_k^{(u,l)}(t) \right|^2}{\sigma_{\text{VLC}}^2}, \tag{40.4}$$

where ε represents the photoelectric-conversion coefficient received by the user, σ^2_{VLC} represents the noise power in the VLC channel, and $p^{(u,l)}_{k,o}$ represents the transmit power of VLC AP_l to user u on subcarrier k.

From Shannon's formula, the effective information-transfer rate between VLC AP_l and user u on subcarrier k of time slot t is

$$\nu^{(u,l)}_k(t) = B_{\mathrm{VLC}} \log_2\left(1 + \xi^{(u,l)}_k(t)\right) \tag{40.5}$$

where B_{VLC} is the subcarrier bandwidth. Because each subcarrier can only be assigned to one user in a single time slot, the following constraints exist for each subcarrier k:

$$\sum_{u \in U} a^{(u,l)}(t) b^{(u,l)}_k(t) \le 1, a^{(u,l)}(t) \in \{0,1\}, b^{(u,l)}_k(t) \in \{0,1\}, \forall l \in L_{\mathrm{VLC}}. \tag{40.6}$$

The transmission rate of VLC AP_l and user u in time slot t is

$$R_{\mathrm{VLC}}(t) = a^{(u,l)}(t)\nu^{(u,l)}(t) = a^{(u,l)}(t) \sum_{k \in K} b^{(u,l)}_k(t)\nu^{(u,l)}_k(t) \tag{40.7}$$

In time slot t, the transmit power of VLC AP_l is

$$P^l_{\mathrm{VLC}}(t) = \sum_{u \in U} a^{(u,l)}(t) P^{(u,l)}(t) = \sum_{u \in U} a^{(u,l)}(t) \sum_{k \in K} b^{(u,l)}_k(t) p^{(u,l)}_k(t) \tag{40.8}$$

To save power, the upper limit of the time-averaged power consumption of VLC AP_l can be set to

$$0 \le \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\left[P^l_{\mathrm{VLC}}(t)\right] < P_{\mathrm{ave,VLC}} \tag{40.9}$$

where $P_{\mathrm{ave,\ VLC}}$ represents the maximum average power consumption on VLC AP_l.

RF channel model

The RF communication path loss is represented by PL [41] and typically takes the following form:

$$\mathrm{PL}[dB] = A \log_{10}\left(d^{(u,l)}\right) + B + C \log_{10}\left(\frac{f_c}{5}\right) + X \tag{40.10}$$

where $d^{(u,\ l)}$ is the distance between the RF AP and the user; f_c is the carrier frequency in GHz; A, B, and C are constants that depend on the propagation model; and X represents the loss caused by shadow effects in different environments.

RF communicates using OFDM, in which the RF AP uses I resource blocks (RBs). In time slot t, $x_i^{(u,0)}(t)$ represents the allocation relationship between user u and the ith RB, if $x_i^{(u,0)}(t) = 1$ and AP_l assigns RB_i to user u; otherwise, $x_i^{(u,0)}(t) = 0$. Because there is only one RF AP in a VLC/RF heterogeneous network, channel interference in the RF system can be ignored.

Assuming that the RF AP transmits power to user u on RB_i of $p_i^{(u,0)}(t)$, the signal-to-noise ratio between user u and the RF AP is

$$\xi_i^{(u,0)}(t) = \frac{p_i^{(u,0)}(t) g_i^{(u,0)}(t)}{\sigma_{\mathrm{RF}}^2}, \tag{40.11}$$

where $g_i^{(u,0)}(t) = 10^{-\mathrm{PL[dB]}/10}$ represents the RF channel power gain and σ_{RF}^2 represents the noise power of the RF channel.

According to Shannon's formula, the information-transmission rate between RF AP subchannel i and user u in time slot t is

$$\nu_i^{(u,0)}(t) = B_{\mathrm{RF}} \log_2\left(1 + \xi_i^{(u,0)}(t)\right), \tag{40.12}$$

where B_{RF} denotes the bandwidth of RB_i. Because each RB can serve only one user in a time slot, the following must be satisfied for each RB_i:

$$\sum_{u \in U} a^{(u,0)}(t) x_i^{(u,0)}(t) \leq 1. \tag{40.13}$$

Therefore, in time slot t, the information-transfer rate between the RF AP and user u is

$$R_{\mathrm{RF}}(t) = a^{(u,0)}(t) \nu^{(u,0)}(t) = a^{(u,0)}(t) \sum_{i \in I} x_i^{(u,0)}(t) \nu_i^{(u,0)}(t). \tag{40.14}$$

The transmit power of the RF AP is derived from the following equation:

$$P_{\mathrm{RF}}^0(t) = \sum_{u \in U} a^{(u,0)}(t) P^{(u,0)}(t) = \sum_{u \in U} a^{(u,0)}(t) \sum_{i \in I} x^{(u,0)}(t) p_i^{(u,0)}(t). \tag{40.15}$$

Considering the power condition, the transmitting power should satisfy

$$P^0 = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} E\left[P_{\mathrm{RF}}^0(t)\right] < P_{\mathrm{ave,RF}}, \tag{40.16}$$

where $P_{\mathrm{ave,RF}}$ is the maximum average power consumption of the RF AP.

Uplink Communication

When the downlink successfully receives a signal, it replies to the sender through an RF transmitter, indicating that the response was successfully received. As RF communication is performed indoors, the Rayleigh fading channel [42] was selected as the upstream RF communication-link channel model. The channel coefficient h of the Rayleigh fading channel follows a complex Gaussian distribution with a mean of 0 and a variance of σ^2, i.e., $h \sim CN(0, \sigma^2)$, where $\sigma^2 = d^{-\alpha}$, d represents the RF communication distance, and α represents the path-fading factor in RF communication. In an indoor environment, the value is usually set as 2–4.

The signal transmitted by the RF transmitter is $\sqrt{P_{\mathrm{RF}}}\chi_{\mathrm{RF}}(t)$, where P_{RF} is the power of the response signal, $\chi_{\mathrm{RF}}(t)$ is the normalized signal, $\chi_{\mathrm{RF}}(t) \in [-1, 1]$, and $E[|\chi_{\mathrm{RF}}(t)|^2] = 1$. The signal at the RF receiver can be expressed as

$$y_{\mathrm{RF}}(t) = \sqrt{P_{\mathrm{RF}}} h \chi_{\mathrm{RF}}(t) + n_{\mathrm{RF}}(t), \tag{40.17}$$

where $n_{\mathrm{RF}}(t)$ is additive white Gaussian noise and follows a complex Gaussian distribution with a mean of zero and variance of σ^2_{RF}.

OFDMA System Model

In this study, a VLC system model based on OFDMA technology was adopted, which aimed to combine optimization objectives with system stability. According to the different M-QAM modulation methods, M-ary QAM can be adaptively applied to different OFDM subcarriers [43]. Different spectral efficiencies (SE) can be achieved for each subcarrier; S_k is used to represent the SE of each subcarrier. The value of S_k depends on the order of the M-QAM. The modulation order of M-QAM is adaptively selected, based on the signal-to-interference-plus-noise ratio (SINR) between the user and the VLC AP. The relationship between S_k and the SINR was determined using the modulation and coding scheme (MCS). These values are listed in Table 40.1.

SINR can be expressed as

$$\mathrm{SINR}(f) = \frac{(\varepsilon P_o H)^2}{N_0 B + \sum\limits_{l \in L} \left(\varepsilon P_o H^{(u,l)}\right)^2}, \tag{40.18}$$

where N_0 represents the noise power spectral density in the VLC channel, B represents the bandwidth of the VLC-baseband modulated signal, H represents the channel gain between the VLC AP providing the service and user u, and $H^{(u,\ l)}$ represents the channel gain between the nonservice VLC AP and user u, which is regarded as interference noise.

The model uses $t = 0, 1, 2, \ldots, T$ as variables by allocating subchannels by time slot, studying the changes in the system as time passes, and reassigning subchannels.

Table 40.1 Relationship between the SINR, modulation, coding rate, and SE [44]

$SINR_{min}$ [dB]	Modulation mode	Encoding rate	SE [bit/Hz]
–	–	–	0
1	QPSK	0.44	0.8770
3	QPSK	0.59	1.1758
5	16QAM	0.37	1.4766
8	16QAM	0.48	1.9141
9	16QAM	0.60	2.4063
11	64QAM	0.45	2.7305
12	64QAM	0.55	3.3223
14	64QAM	0.65	3.9023
16	64QAM	0.75	4.5234
18	64QAM	0.85	5.1152
20	64QAM	0.93	5.5547

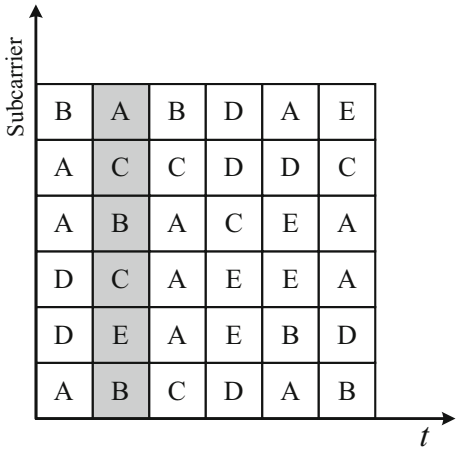


Fig. 40.3 OFDMA system resource-allocation model

This method fundamentally improves the frequency of system resource allocation, such that the system can quickly make optimal allocations as time passes. In addition, we used the Lyapunov optimization method for this optimization process to ensure that the system remained stable during changes within $t = 0, 1, 2, \ldots, T$.

Figure 40.3 shows the OFDMA system resource-allocation model. The dark part represents the resource-allocation scale; for each RB, the smallest block of indivisible resources is called the resource unit (RU). The RU allocates resources in each time slot, and each RU allocation process is independent. This allocation process differs from one-dimensional resource allocation in that in each time slot, the resource-allocation problem in OFDMA is degraded from two dimensional to one dimensional by setting the integer restriction to $b_k^u(t) \in \{0, 1\}$, and the integer allocation is reduced to a 0–1 integer-allocation problem.

In VLC systems, it is generally believed that the channel is time-flat and that there is no fading during communication. Therefore, the channel-state information is constant for each time slot. Suppose that the VLC AP providing the service and subscriber-device subcarrier k can achieve the rate represented by the following equation:

$$\nu_{k,u}(t) = 2BS_k/K. \tag{40.19}$$

The transmission rate between the user on subchannel k and the VLC AP providing the service is

$$R_{k,u}(t) = b_k^u(t)\nu_{k,u}(t). \tag{40.20}$$

The transfer rate provided by the VLC AP to user u is

$$R_u(t) = \sum_{k \in K} R_{k,u}(t). \tag{40.21}$$

Because a subcarrier can serve only one user per time slot, its distribution coefficient must satisfy the following constraints:

$$\sum_{u \in U} b_k^u(t) = 1, b_k^u(t) \in \{0, 1\}, \forall k \in K. \tag{40.22}$$

Queue Model

When the system begins transmitting data, let $Q_u(t)$ represent the length of the queue that the system maintains for user u at the beginning of time slot t. $D_u(t)$ represents the amount of data that user u sends at time slot t, which is independently and identically distributed across different time slots [45]. Assuming that the average data-arrival rate of user u in time slot t is $\gamma_u(t)$, the following conditions are met for $D_u(t)$:

$$E[D_u(t)] = \gamma_u(t), \forall u \in U, \tag{40.23}$$

where γ_u is a bounded variable that guarantees the stability of the system. According to the above analysis, $Q_u(t)$ can be expressed as

$$Q_u(t+1) = [Q_u(t) - R_u(t), 0]^+ + D_u(t), \tag{40.24}$$

where $R_u(t)$ represents the amount of data left and $[x, 0]^+ = \max[x, 0]$.

For each user u, the randomness of $D_u(t)$ and the time-varying nature of $R_u(t)$ are provided to make the queuing process random. If we only consider maximizing the system-transfer rate and ignore the impact of the queue, data will accumulate in the

queue during transmission, resulting in the system operating abnormally. Therefore, a queue-stability model must be established; it is generally defined as

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}E[Q_u(t)]<\infty. \tag{40.25}$$

If all queues in the system are stable, the VLC communication system will also be stable. This means that the average departure rate in the queue is greater than or equal to the average input rate. Once this condition is met, all data in the queue are sent to the corresponding user.

Optimization Theory

The goal of this study was to achieve the maximum transfer rate, while keeping the system stable. Within time slot t, the transmission rate of the entire system can be expressed as

$$R(t)=\sum_{u\in U}R_u(t)=\sum_{u\in U}\sum_{l\in L}R^{(u,l)}(t). \tag{40.26}$$

To achieve the optimization goal of maximizing the average system transmission rate, the above function was designed as a time-averaged expression. Therefore, the stochastic optimization problem was formulated as follows:

$$\begin{cases}
\max\ \lim\limits_{T\to\infty}\dfrac{1}{T}\sum\limits_{t=0}^{T-1}E\left[R(t)\right] & \\
\text{s.t}\ \lim\limits_{T\to\infty}\dfrac{1}{T}\sum\limits_{t=0}^{T-1}E\left[Q_u(t)\right]<\infty,\ \forall u\in U & \text{(a)}\\
\sum\limits_{l\in L}a^{(u,l)}(t)=1,\ a^{(u,l)}(t)\in\{0,1\},\ \forall u\in U,\ \forall l\in L & \text{(b)}\\
\sum\limits_{u\in U}a^{(u,0)}(t)x_i^{(u,0)}(t)\le 1,\ a^{(u,l)}(t)\in\{0,1\},\ x_i^{(u,0)}(t)\in\{0,1\},\ \forall u\in U & \text{(c)}\\
0\le\lim\limits_{T\to\infty}\dfrac{1}{T}\sum\limits_{t=0}^{T-1}E\left[P_l(t)\right]<P_{l,ave},\ P_{l,ave}=\begin{cases}P_{ave,RF},\ \text{if } l=0\\ P_{ave,VLC},\ \text{if } l=1,\cdots L\end{cases} & \text{(d)}\\
\sum\limits_{u\in U}a^{(u,l)}(t)b_k^{(u,l)}(t)\le 1,\ a^{(u,l)}(t)\in\{0,1\},\ b_k^{(u,l)}(t)\in\{0,1\},\ \forall l\in L_{VLC} & \text{(e)}\\
\sum\limits_{u\in U}b_k^{u}(t)=1,\ \forall k\in K & \text{(f)}\\
P_i^{(u,0)}(t)\ge 0,\ P_k^{(u,l)}(t)\ge 0,\ \forall l\in L_{VLC},\forall u\in U & \text{(g)}
\end{cases} \tag{40.27}$$

where (a) is the system stability requirement; (b) ensures that in each time slot, the user can only be served by one AP; (c) ensures that for the RF AP, only one user can be served on each resource block; (d) represents the average power-budget constraint of each AP; (e) ensures that each subchannel of the VLC AP can only serve one user; (f) ensures that for one VLC AP, each subcarrier can only be assigned to one user at a time; and (g) means that the transmit power of all APs is nonnegative and meets the actual conditions.

The above problem has long-term constraints. To effectively handle long-term time-average power constraints, virtual queues in Lyapunov optimization can be used to convert (d) into queue-stability constraints [46]. Virtual queues are defined as $Z(t) \cong \{Z_1(t),\ldots, Z_l(t),\ldots, Z_L(t)\}$. $Z_l(t)$ evolves as follows:

$$Z_l(t+1) = \max\{Z_l(t) - P_{l,\text{ave}}, 0\} + P_l(t). \tag{40.28}$$

It should be noted that there is no actual queue data in queue $Z_l(t)$; it is only proposed to satisfy constraint (d). If virtual queue $Z_l(t)$ is average-rate-stable, constraint (d) automatically holds [47].

Suppose the order vector of the queue length in the system is $\Theta(t) = [Q(t), Z(t)]$. The Lyapunov function is defined as a measure of the total queue length for each time slot t and is denoted by the following equation:

$$\mathrm{L}(\Theta(t)) = \frac{1}{2}\sum_{\substack{u \in U \\ l \in L}} Q_u^2(t) + \frac{1}{2}\sum_{l \in L} Z_l^2(t). \tag{40.29}$$

In each time slot, the conditional Lyapunov drift is defined as

$$\Delta(\Theta(t)) \triangleq E[\mathrm{L}(\Theta(t+1)) - \mathrm{L}(\Theta(t))|\Theta(t)]. \tag{40.30}$$

In each time slot, the Lyapunov function controls the final value of the Lyapunov drift by adjusting the final queue length. The purpose of this study is to maximize the average system-transfer rate. The drift-penalty function term is defined as

$$\Delta(\Theta(t)) - VE[R(t)|\Theta(t)], \tag{40.31}$$

where V is a nonnegative constant parameter used to balance the drift with a reward. When V increases, the system assigns a higher priority to maximize the transmission rate; that is, it sacrifices the queue length. When the V value decreases, the system prioritizes the queue length to maintain the system stability.

For each time period, for a given $V \geq 0$, the upper bound of the drift-penalty function for any possible $\Theta(t)$ is

$$\begin{aligned}\Delta(\Theta(t)) - VE[R(t)|\Theta(t)] &= E[\mathrm{L}(\Theta(t+1)) - \mathrm{L}(\Theta(t))|\Theta(t)] - V\sum_{\substack{u\in U\\ l\in L}} E\left[R^{(u,l)}(t)|\Theta(t)\right] \\ &\le C + \sum_{\substack{u\in U\\ l\in L}} E\left[Q_u(t)\left(D_u(t) - a^{(u,l)}(t)\nu^{(u,l)}(t)\right)|\Theta(t)\right] \\ &\quad + \sum_{l\in L} E[Z_l(t)(P_l(t) - P_{l,\mathrm{ave}})|\Theta(t)],\end{aligned} \tag{40.32}$$

where C is a constant satisfying the following conditions:

$$C \ge \frac{1}{2}\sum_{u\in U} E\left[D_u^2(t) + R_u^2(t)|\Theta(t)\right] + \frac{1}{2}\sum_{l\in L} E\left[P_l^2(t) + (P_{l,\mathrm{ave}})^2|\Theta(t)\right]. \tag{40.33}$$

According to the Lyapunov optimization method [48], the network utility is optimized by minimizing the drift-penalty function term to minimize the queue length. Therefore, Eq. (40.28) can be transformed into minimizing the upper bound of the Lyapunov drift-penalty function in each time slot, which is expressed as

$$\begin{cases} \min \sum\limits_{\substack{u\in U\\ l\in L}} Q_u(t)\left(D_u(t) - a^{(u,l)}(t)\nu^{(u,l)}(t)\right) + \sum\limits_{l\in L} Z_l(t)(P_l(t) - P_{l,\mathrm{ave}}) - V\sum\limits_{\substack{u\in U\\ l\in L}} a^{(u,l)}(t)\nu^{(u,l)}(t) \\ \text{s.t (27c) (27d) (27f) (27h).} \end{cases} \tag{40.34}$$

The reformulated optimization problem, Eq. (40.34), is a constrained mixed-integer programming problem, where $a^{(u,\,l)}(t)$, $x_i^{(u,0)}(t)$, and $b_k^{(u,l)}(t)$ take integer values, and $p_i^{(u,0)}(t)$ and $p_k^{(u,l)}(t)$ take continuous values. Therefore, optimization problem (40.34) is decomposed into three relatively independent subproblems, and a corresponding optimization algorithm is proposed.

AP Allocation

The AP allocation variables $a^{(u,\,l)}(t)$ in the system can be obtained by solving the following subproblems:

$$\begin{cases} \min \sum\limits_{\substack{u\in U\\ l\in L}} Q_u(t)\left(D_u(t) - a^{(u,l)}(t)\nu^{(u,l)}(t)\right) - V\sum\limits_{\substack{u\in U\\ l\in L}} a^{(u,l)}(t)\nu^{(u,l)}(t) \\ \text{s.t} \sum\limits_{l\in L} a^{(u,l)}(t) = 1, a^{(u,l)}(t) \in \{0,1\}, \forall u \in U, \forall l \in L. \end{cases} \tag{40.35}$$

From (40.35), it can be observed that the network selection between different users is independent. Therefore, problem (40.35) can be decomposed into U subproblems, where each subproblem is

$$\begin{cases} \max \sum\limits_{l \in L} a^{(u,l)}(t) e^{(u,l)}(t) \\ \text{s.t} \sum\limits_{l \in L} a^{(u,l)}(t) = 1, a^{(u,l)}(t) \in \{0,1\}, \forall u \in U, \forall l \in L, \end{cases} \tag{40.36}$$

where $e^{(u,\ l)}(t) = (Q_u(t) + V)\, v^{(u,\ l)}$. As (40.36) is a linear-optimization problem, its solution is as follows:

$$a^{(u,l)}(t) = \begin{cases} 1, l = \arg\max\limits_{l'} \left\{ e^{(u,l')}(t), 0 < l' < L \right\}, \\ 0, \text{otherwise}. \end{cases} \tag{40.37}$$

Under the above network-selection strategy, $e^{(u,l)}(t)$ can be regarded as the congestion cost of the AP serving the user within time slot t. Then, Eq. (40.37) ensures that the user can connect to the AP with the lowest congestion cost.

VLC AP Joint Subcarrier and Power Allocation

According to the access-point allocation variable $a(t)$, obtained from Eq. (40.35), the joint subcarrier and power-allocation subproblems in the VLC system can be continuously solved.

$$\begin{cases} \max \sum\limits_{\substack{u \in U \\ l \in L_{\text{VLC}}}} a^{(u,l)}(t) \nu^{(u,l)}(t) (Q_u(t) + V) - Z_l(t)(P_l(t) - P_{\text{ave,VLC}}(t)) \\ \text{s.t} \sum\limits_{u \in U} a^{(u,l)}(t) b_k^{(u,l)}(t) \le 1, a^{(u,l)}(t) \in \{0,1\}, b_k^{(u,l)}(t) \in \{0,1\}, \forall u \in U, \\ p_k^{(u,l)}(t) \ge 0, \forall u \in U. \end{cases} \tag{40.38}$$

Because the selection of subcarriers is independent among different VLC APs, problem (40.38) can be decomposed into $|L_{\text{VLC}}|$ subproblems, and each subproblem is expressed as

$$\begin{cases} \max \sum\limits_{u \in U} B_{\text{VLC}} \log_2 \left(1 + \dfrac{p_k^{(u,l)}(t) \left| \varepsilon h_k^{(u,l)}(t) \right|^2}{\sigma_{\text{VLC}}^2} \right) (Q_u(t) + V) - Z_l(t)(P_l(t) - P_{\text{ave,VLC}}(t)) \\ \text{s.t} \sum\limits_{u \in U} a^{(u,l)}(t) b_k^{(u,l)}(t) \le 1, a^{(u,l)}(t) \in \{0,1\}, b_k^{(u,l)}(t) \in \{0,1\}, \forall u \in U_l, \\ p_k^{(u,l)}(t) \ge 0, \forall u \in U_l, \end{cases} \tag{40.39}$$

where U_l denotes the set of users connected to the VLC AP. The VLC joint subcarrier and power-allocation problem is a hybrid combinatorial optimization problem. To solve this problem, the Lagrange relaxation method can be used to transform the original mixed-integer-programming problem into a convex-optimization problem. Relaxing $b_k^{(u,l)}(t) \in \{0,1\}$ to $\tilde{b}_k^{(u,l)}(t) \in [0,1]$, the optimization problem is reformulated as

$$\begin{cases} \max \sum\limits_{k \in K} \sum\limits_{u \in U_l} \tilde{b}_k^{(u,l)}(t)(Q_u(t)+V)B_{\mathrm{VLC}}\log_2\left(1+\dfrac{p_k^{(u,l)}(t)\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\tilde{b}_k^{(u,l)}(t)\sigma_{\mathrm{VLC}}^2}\right) \\ \quad - Z_l(t)\left(\sum\limits_{k \in K}\sum\limits_{u \in U} p_k^{(u,l)}(t) - P_{\mathrm{ave,VLC}}(t)\right) \\ \text{s.t} \sum\limits_{u \in U_l} \tilde{b}_k^{(u,l)}(t) \le 1, \forall u \in U_l, \\ \quad 0 \le \tilde{b}_k^{(u,l)}(t) \le 1, \forall u \in U_l, \\ \quad p_k^{(u,l)}(t) \ge 0, \forall u \in U_l. \end{cases} \tag{40.40}$$

As $\sum\limits_{k \in K}\sum\limits_{u \in U_l} \tilde{b}_k^{(u,l)}(t)(Q_u(t)-V)B_{\mathrm{VLC}}\log_2\left(1+\dfrac{p_k^{(u,l)}(t)\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\tilde{b}_k^{(u,l)}(t)\sigma_{\mathrm{VLC}}^2}\right)$ is a convex function of $\tilde{b}_k^{(u,l)}(t)$ and $p_k^{(u,l)}(t)$, it is considered a linear function of $\log_2\left(1+\frac{p_k^{(u,l)}(t)\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\sigma_{\mathrm{VLC}}^2}\right)$. Therefore, the given optimization problem is a $K \times |U_l|$ linear combination of convex functions.

Let $\mathrm{L}(P) = \sum\limits_{k \in K}\sum\limits_{u \in U_l} \tilde{b}_k^{(u,l)}(t)(Q_u(t)+V)B_{\mathrm{VLC}}\log_2\left(1+\frac{p_k^{(u,l)}(t)\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\tilde{b}_k^{(u,l)}(t)\sigma_{\mathrm{VLC}}^2}\right) - Z_l(t)\sum\limits_{k \in K}\sum\limits_{u \in U} p_k^{(u,l)}(t)$ be the objective function and determine the partial derivative of $p_k^{(u,l)}(t)$ for $\mathrm{L}(P)$.

$$\frac{\partial \mathrm{L}(P)}{\partial p_k^{(u,l)}(t)} = \frac{\tilde{b}_k^{(u,l)}(t)(Q_u(t)+V)B_{\mathrm{VLC}}\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\left(\tilde{b}_k^{(u,l)}(t)\sigma_{\mathrm{VLC}}^2 + p_k^{(u,l)}(t)\left|\varepsilon h_k^{(u,l)}(t)\right|^2\right)\ln 2} - Z_l(t) \tag{40.41}$$

According to the Karush–Kuhn–Tucker (KKT) condition, the system can achieve optimal power allocation when the following conditions are satisfied:

$$\begin{aligned} &\frac{\partial\left[\mathrm{L}(P) - \left(\sum\limits_{k \in K}\sum\limits_{u \in U_l} \mu_k^u p_k^{(u,l)}(t)\right)\right]}{\partial p_k^{(u,l)}(t)} \\ &= \frac{\tilde{b}_k^{(u,l)}(t)(Q_u(t)+V)B_{\mathrm{VLC}}\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\left(\tilde{b}_k^{(u,l)}(t)\sigma_{\mathrm{VLC}}^2 + p_k^{(u,l)}(t)\left|\varepsilon h_k^{(u,l)}(t)\right|^2\right)\ln 2} \end{aligned} \tag{40.42}$$

$$-Z_l(t) - \mu_k^u = 0 \tag{40.43}$$

$$\mu_k^u p_k^{(u,l)}(t) = 0, \forall k, u \tag{40.44}$$

$$\mu_k^u \geq 0, \forall k, u, \tag{40.45}$$

where μ_k^u is a multiplier used to construct the Lagrange function.

To ensure that (40.44) holds true, we use $\mu_k^u = 0$ or $p_k^{(u,l)}(t) = 0$. The partial derivative of (40.42) is 0, and the relationship between $p_k^{(u,l)^*}(t)$ and $\tilde{b}_k^{(u,l)}(t)$ is obtained.

$$p_k^{(u,l)^*}(t) = \left[\frac{(Q_u(t) + V)B_{\mathrm{VLC}}}{Z_l(t)\ln 2} - \frac{\sigma_{\mathrm{VLC}}^2}{\left|\varepsilon h_k^{(u,l)}(t)\right|^2} \right]^+ \tilde{b}_k^{(u,l)}(t). \tag{40.46}$$

The power-allocation result is used to obtain subcarrier allocation. By substituting Eq. (40.46) into Equation $\mathrm{L}(P)$, we obtain

$$\mathrm{L}(P) = \sum_{k \in K} \sum_{u \in U} L_k^u(P) \tilde{b}_k^{(u,l)}(t), \tag{40.47}$$

where

$$L_k^u(P) = (Q_u(t) + V)B_{\mathrm{VLC}} \log_2 \left(1 + \left[\frac{(Q_u(t) + V)B_{\mathrm{VLC}}}{Z_l(t)\ln 2} - \frac{\sigma_{\mathrm{VLC}}^2}{\left|\varepsilon h_k^{(u,l)}(t)\right|^2} \right]^+ \frac{\left|\varepsilon h_k^{(u,l)}(t)\right|^2}{\sigma_{\mathrm{VLC}}^2} \right)$$
$$-Z_l(t) \left[\frac{(Q_u(t) + V)B_{\mathrm{VLC}}}{Z_l(t)\ln 2} - \frac{\sigma_{\mathrm{VLC}}^2}{\left|\varepsilon h_k^{(u,l)}(t)\right|^2} \right]^+. \tag{40.48}$$

Equation (40.47) is a linear-assignment problem that can be decomposed into k subproblems. Each subproblem is represented by the following equation:

$$\mathrm{L}_k(P) = \sum_{u \in U} L_k^u(P) \tilde{b}_k^{(u,l)}(t). \tag{40.49}$$

For subcarrier k, the optimal allocation obtains the user with the maximum $L_k^u(P)$, and the solution can be obtained:

$$\tilde{b}_k^{(u,l)^*}(t) = \begin{cases} 1, u = \underset{u'}{\mathrm{argmax}} \left(L_k^{u'}(P), u' \in U_l \right), L_k^{u'}(P) > 0, \\ 0, \text{otherwise.} \end{cases} \tag{40.50}$$

RF AP Joint Resource Block and Power Allocation

In RF communication systems, the RF AP joint resource block and power-allocation subproblem can be expressed as

$$\begin{cases} \max \sum\limits_{u \in U_0} a^{(u,0)}(t)(Q_u(t)+V) - Z_0(t)(P_0(t) - P_{\text{ave,RF}}(t)) \\ \text{s.t} \sum\limits_{u \in U_0} x_i^{(u,0)}(t) \le 1, a^{(u,0)}(t) \in \{0,1\}, x_i^{(u,0)}(t) \in \{0,1\}, \forall u \in U_0 . \\ p_i^{(u,0)}(t) \ge 0, \forall u \in U_0 \end{cases} \quad (40.51)$$

It can be seen from the above equations that the joint RB and power-allocation problems are independent of each other on different time slots t. Therefore, the objective function of problem (40.51) is a function of discrete variable $x_i^{(u,0)}(t)$ and continuous variable $p_i^{(u,0)}(t)$; thus, the problem is also a mixed-integer nonlinear convex-optimization problem, which can be solved by the Lagrange relaxation method.

Firstly, $x_i^{(u,0)}(t)$ is relaxed to the continuous interval [0, 1], and a new variable $\tilde{x}_i^{(u,0)}(t)$ is defined. Problem (40.51) can be rewritten as:

$$\begin{cases} \max \sum\limits_{i \in I} \sum\limits_{u \in U_0} \tilde{x}_i^{(u,0)}(t)(Q_u(t)+V) B_{\text{RF}} \log_2 \left(1 + \dfrac{p_i^{(u,0)}(t) g_i^{(u,0)}(t)}{\tilde{x}_i^{(u,0)}(t)\sigma_{\text{RF}}^2}\right) - Z_0(t) \sum\limits_{i \in I} \sum\limits_{u \in U_0} p_i^{(u,0)}(t) \\ \text{s.t} \sum\limits_{u \in U_0} \tilde{x}_i^{(u,0)}(t) \le 1, \tilde{x}_i^{(u,0)}(t) \in [0,1], \forall u \in U_0 \\ p_i^{(u,0)}(t) \ge 0, \forall u \in U_0 \end{cases} . \quad (40.52)$$

Let the objective function be

$$\mathrm{L}(P) = \sum_{i \in I} \sum_{u \in U_0} \tilde{x}_i^{(u,0)}(t)(Q_u(t)+V) B_{\text{RF}} \log_2 \left(1 + \frac{p_i^{(u,0)}(t) g_i^{(u,0)}(t)}{\tilde{x}_i^{(u,0)}(t)\sigma_{\text{RF}}^2}\right) - Z_0(t) \sum_{i \in I} \sum_{u \in U_0} p_i^{(u,0)}(t),$$

and calculate the partial derivative of $p_i^{(u,0)}(t)$:

$$\frac{\partial \mathrm{L}(P)}{\partial p_i^{(u,0)}(t)} = \frac{\tilde{x}_i^{(u,0)}(t)(Q_u(t)+V) B_{\text{RF}} g_i^{(u,0)}(t)}{\left(\tilde{x}_i^{(u,0)}(t)\sigma_{\text{RF}}^2 + p_i^{(u,0)}(t) g_i^{(u,0)}(t)\right) \ln 2} - Z_0(t). \quad (40.53)$$

According to the KKT condition, the system achieves the optimal power allocation when the following conditions are satisfied:

$$\frac{\partial \left[\mathrm{L}(P) - \left(\sum\limits_{i \in I} \sum\limits_{u \in U_0} \mu_i^u p_i^{(u,0)}(t) \right) \right]}{\partial p_i^{(u,0)}(t)} = \frac{\tilde{x}_i^{(u,0)}(t)(Q_u(t)+V) B_{\text{RF}} g_i^{(u,0)}(t)}{\left(\tilde{x}_i^{(u,0)}(t)\sigma_{\text{RF}}^2 + p_i^{(u,0)}(t) g_i^{(u,0)}(t)\right) \ln 2} \quad (40.54)$$

$$-Z_0(t) - \mu_i^u = 0 \tag{40.55}$$

$$\mu_i^u p_i^{(u,0)}(t) = 0, \forall i, u \tag{40.56}$$

$$\mu_i^u \geq 0, \forall i, u, \tag{40.57}$$

where μ_i^u is a multiplier used to construct the Lagrange function. To ensure that Eq. (40.56) holds, then $\mu_i^u = 0$ or $p_i^{(u,l)}(t) = 0$. When the partial derivative of Eq. (40.54) is zero, the relationship between $p_i^{(u,l)^*}(t)$ and $\tilde{x}_i^{(u,l)}(t)$ can be obtained:

$$p_i^{(u,0)^*}(t) = \left[\frac{(Q_u(t)+V)B_{\mathrm{RF}}}{Z_0(t)\ln 2} - \frac{\sigma_{RF}^2}{g_i^{(u,0)}(t)}\right]^+ \tilde{x}_i^{(u,0)}(t). \tag{40.58}$$

The results of the power allocation are used to allocate the RF resource blocks. Substituting Eq. (40.56) into $\mathrm{L}(P)$, we obtain

$$\mathrm{L}(P) = \sum_{i \in I} \sum_{u \in U_0} L_i^u(P)\tilde{x}_i^{(u,0)}(t), \tag{40.59}$$

where

$$\begin{aligned} L_i^u(P) = {} & a_i^{(u,0)}(t)(Q_u(t)+V)B_{\mathrm{RF}}\log_2\left(1 + \left[\frac{a_i^{(u,0)}(t)\tilde{x}_i^{(u,0)}(t)(Q_u(t)+V)B_{\mathrm{RF}}}{Z_0(t)\ln 2} - \frac{B_{\mathrm{RF}}\sigma_{\mathrm{RF}}^2}{g_i^{(u,0)}(t)}\right]^+ \frac{g_i^{(u,0)}(t)}{B_{\mathrm{RF}}\sigma_{\mathrm{RF}}^2}\right) \\ & - Z_0(t)\left[\frac{a_i^{(u,0)}(t)\tilde{x}_i^{(u,0)}(t)(Q_u(t)+V)B_{\mathrm{RF}}}{Z_0(t)\ln 2} - \frac{B_{\mathrm{RF}}\sigma_{\mathrm{RF}}^2}{g_i^{(u,0)}(t)}\right]^+. \end{aligned} \tag{40.60}$$

Problem (40.59) is also a linear-assignment problem. Therefore, it can be decomposed into i subproblems, each of which is expressed as

$$\mathrm{L}_i(P) = \sum_{u \in U} L_i^u(P)\tilde{x}_i^{(u,0)}(t). \tag{40.61}$$

The objective of this problem is to maximize the $\mathrm{L}_i(P)$ of all subcarriers. Therefore, for all subcarriers i, the optimal allocation involves selecting the user with the largest $\mathrm{L}_i(P)$.

$$\tilde{x}_i^{(u,l)^*}(t) = \begin{cases} 1, u = \underset{u'}{\mathrm{argmax}}\left(L_i^{u'}(P), u' \in U\right) \\ 0, \text{otherwise.} \end{cases} \tag{40.62}$$

Algorithm Simulation Results Analysis

In this section, we analyze the simulation results of the Lyapunov optimization algorithm and verify the performance of the proposed indoor VLC/RF system resource-allocation scheme using OFDMA technology. Assuming that the room size is 8 m × 8 m × 3 m, as shown in Fig. 40.1, four VLC APs are installed on the ceiling, which are located at (3, 2, 3), (5, 2, 3), (3, 5, 3) and (5, 5, 3), respectively, and each AP has 25 subcarriers. The RF AP is located at (0, 0, 0), and it has four RBs. The other key simulation parameters of the system are listed in Table 40.2.

To simplify the simulation, we consider the following:

1. The VLC AP and RF AP have a normalized spectrum bandwidth.
2. The traffic arrival for each user device follows the Poisson distribution.
3. The average traffic arrival rate is expressed in γ.
4. The traffic arrival rate and RF channel conditions remain constant within a time slot and change on the time-slot boundary.
5. Specific γ values are given in the simulation.

The spectral bandwidth is normalized and the unit of transmitted data volume is set to one package. Therefore, assuming that the average communication rate of the VLC AP subcarrier in the system is five packages, and the average communication rate of a single RF AP RB is three packages, the average traffic arrival rate is limited to less than five packages.

Figure 40.4a shows the relationship between the penalty parameter V and the average queue length Q under different average traffic arrival rates γ. It can be observed that the Q value in the system increases with the increase of V, and as the average traffic arrival rate γ increases, the average queue length Q also increases.

Figure 40.4b shows the change of penalty parameter V and average transmission rate R under different average traffic arrival rates γ. It can be observed that the average transfer rate R increases with an increase in V because when V increases, the algorithm's optimization of the average transmission rate increases.

Table 40.2 Simulation parameters

Parameter	Meaning	Value
L × W × H	Room size	8 m × 8 m × 3 m
$\Phi_{1/2}$	Power half angle	70°
FOV	Field of view	60°
n	Number of LEDs	5
P_{ave}	LED transmit power	4 W
A_p	Effective PD receiving area	1 cm^2
T_l	Optical-filter gain	1
G_l	Concentrator gain	1
ε	Photoelectric conversion factor	0.52 A/W
n_{re}	PD refractive index	1.5

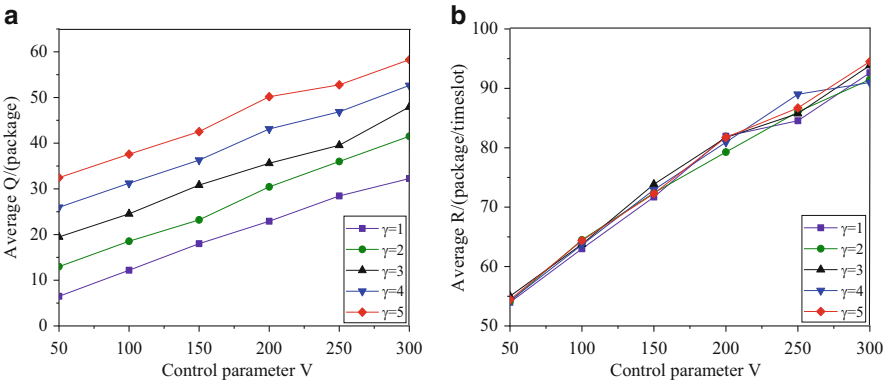


Fig. 40.4 (**a**) Relationship between average queue length Q and control parameter V. (**b**) Relationship between average transmission rate R and control parameter V

In addition, as the average traffic arrival rate γ increases, the growth rate of the average transmission rate decreases, because when γ is small, the amount of data transmitted by the system can meet the amount of arriving data. When γ is large, the system rejects additional packets to ensure data-queue stability.

According to Little's theorem [49], because the average transmission delay is proportional to the average queue length, the average queue length can be used to describe the average delay. In the Lyapunov optimization process, selecting the appropriate control parameter V can balance the relationship between the average transmission rate and average transmission delay.

In addition, the instantaneous change in the system was simulated. Assuming that the user's position changes in each time slot, the instantaneous change of the user in 30 time slots is given.

Figures 40.5a–c depict the instantaneous transfer rate and queue-length change for the four users in the system, when the penalty parameter V was 50, 150, and 300. It can be observed that for each user, as their queue length increases, so does their data-transfer rate. This is because the Lyapunov optimization algorithm increases the system transfer rate to reduce the buffer queue, thereby ensuring the stability of the entire system.

Another conclusion was that the queue length in the system was longer when $V = 300$ than when $V = 50$ or 150.

Figure 40.6 shows the changes in the average queue length Q and average transfer rate R with the number of users in the system using different existing algorithms. The Lyapunov optimization algorithm was compared with the minimal distance resource-allocation (MD-RA) and best channel state information resource-allocation (BCSI-RA) algorithms [50].

From Fig. 40.6a, it can be seen that with the increase in the number of users, the average queue length of the system gradually increases. Compared with the MD-RA algorithm and BCSI-RA algorithm, the queue length under the Lyapunov

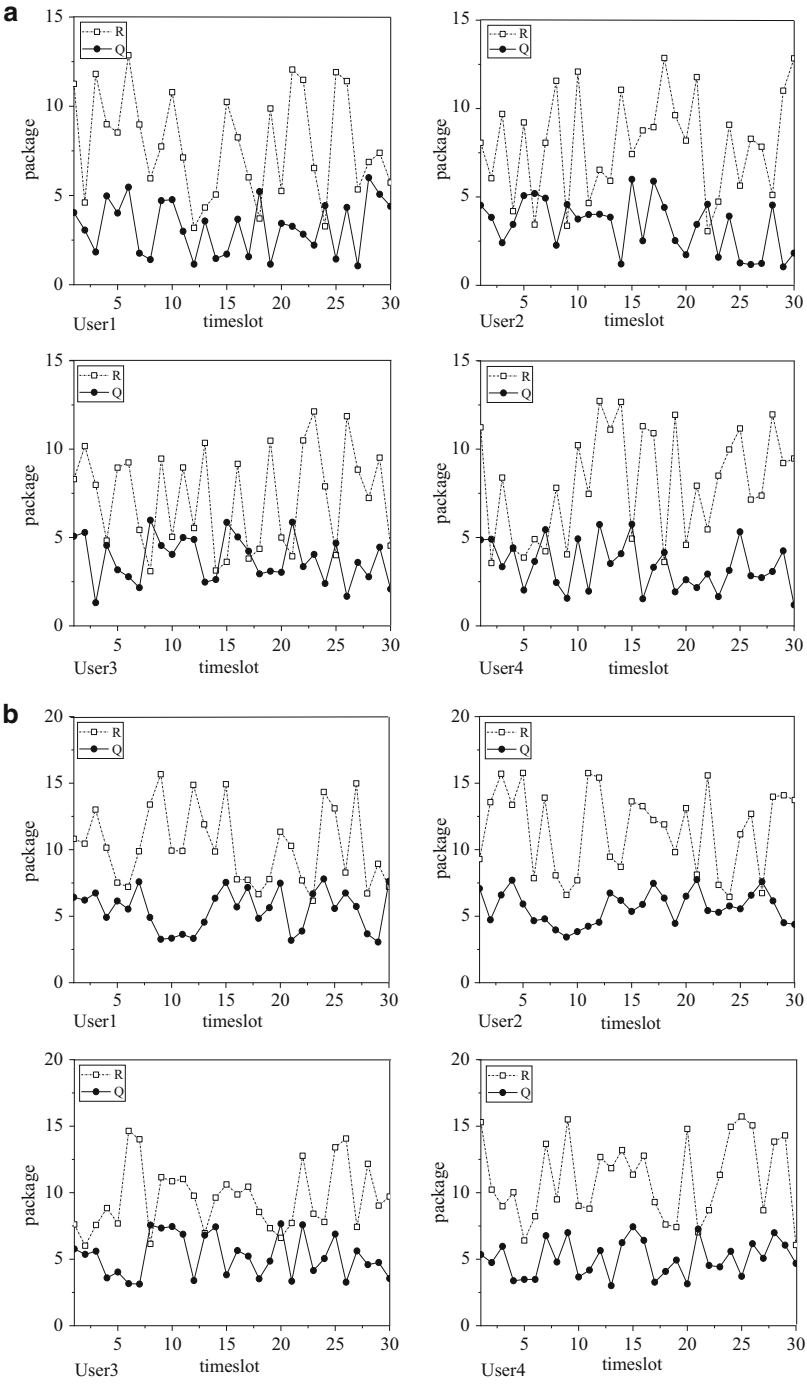


Fig. 40.5 Change in the instantaneous transmission rate and queue length of each user. (**a**) $V = 50$. (**b**) $V = 150$. (**c**) $V = 300$

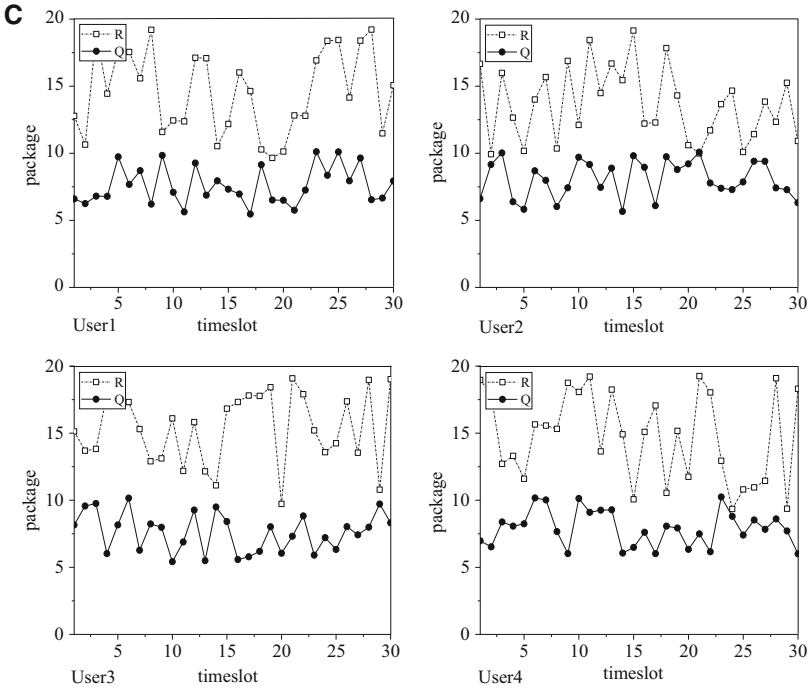


Fig. 40.5 (continued)

optimization algorithm is smaller and the growth rate is slower, indicating that the system has better stability when using the Lyapunov algorithm.

It can be seen from Fig. 40.6b that as the number of users in the system increases, the average transmission rate also increases, and the average transmission rate under the Lyapunov optimization algorithm is significantly higher than those of the MD-RA and BCSI-RA algorithms. The average transmission rate increases significantly when the number of users is greater than six.

Figure 40.7 shows the average queue length versus the average transfer rate as a function of the number of users, considering the interrupt probability, and compares the system performance using the three algorithms.

Suppose p represents the probability of LOS availability in the VLC system and a random number is generated between 0 and 1. If this number is greater than p, the LOS of the VLC AP is available; otherwise, the LOS is not available. The study showed that the Lyapunov optimization algorithm can achieve higher transmission rates while ensuring that the average queue length of the system is low; that is, the algorithm provides better system performance.

As shown in Fig. 40.8, as the penalty parameter V increases, the average queue length Q and average transmission rate R in the uplink communication link for different numbers of users also increase. This is because increasing the penalty

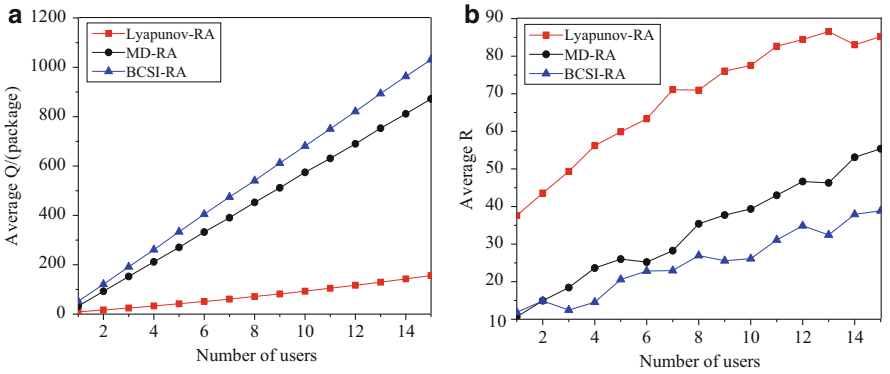


Fig. 40.6 (**a**) Relationship between the average queue length Q and number of users under the three algorithms. (**b**) Relationship between the average transmission rate R and number of users under the three algorithms

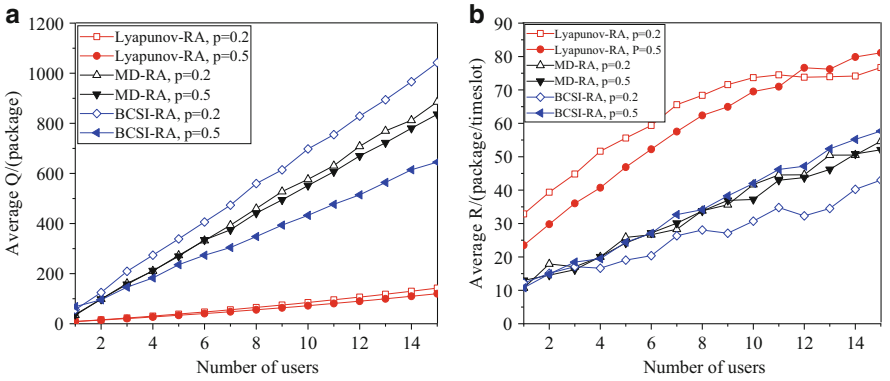


Fig. 40.7 (**a**) Relationship between the average queue length Q, interrupt probability, and number of users for the three algorithms. (**b**) Relationship between the average transmission rate R, interrupt probability, and number of users for the three algorithms

parameter causes the system to allocate more resources to the user, thereby optimizing the system performance.

In addition, as shown in Fig. 40.8a, as the number of users increases, the queue backlog in the system increases at an accelerated pace. An increase in the number of users leads to a decrease in the uplink communication rate, which affects the performance of the entire system. To balance the relationship between the queue length and transmission rate, an appropriate penalty parameter V can be selected to stabilize the system communication.

Figure 40.8b shows that the average transfer rate gradually increases as the number of users increases. Changes in both Q and R indicate a decrease in the communication rate of the upstream RF link. This is because, with limited spectrum

resources and base station processing power, users must share resources, resulting in network congestion and transmission delays, which in turn reduces the communication rate of the entire system.

Figure 40.9 shows the average queue length Q and average transmission rate R in the OFDMA-VLC/RF and VLC/RF systems as functions of the number of users. As shown in Fig. 40.9a, the queue-extrusion speed in the VLC/RF system using OFDMA technology is significantly slower than that in the VLC/RF system without OFDMA technology. The Q-value change in the OFDMA-VLC/RF system is relatively stable, indicating that the system is highly stable.

As shown in Fig. 40.9b, as the number of users increases, the average system transmission rate gradually increases; when the number of users increases to 15, the

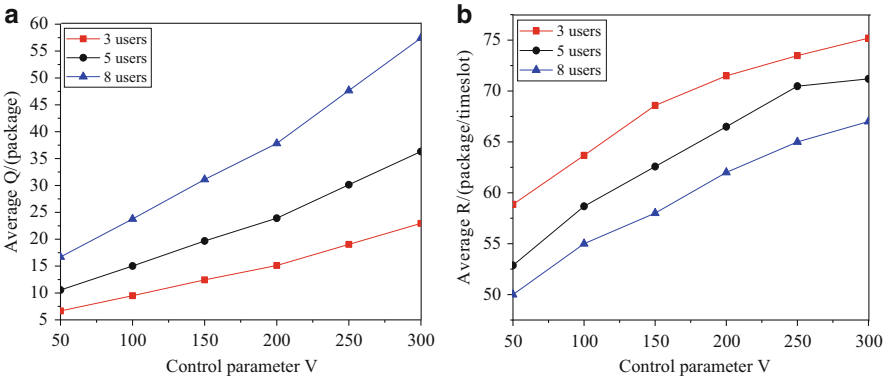


Fig. 40.8 (**a**) Relationship between the average queue length Q and penalty parameter V for different numbers of users in the uplink RF link. (**b**) Relationship between the average transmission rate R and penalty parameter V of the uplink RF link for different numbers of users

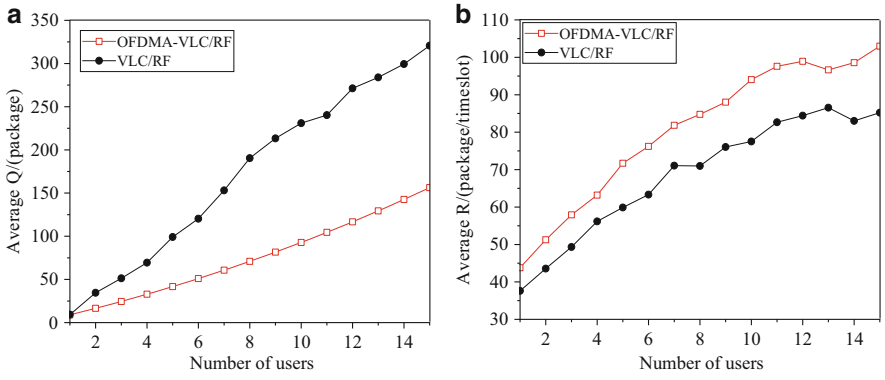


Fig. 40.9 (**a**) Comparison of the average queue length Q of the system. (**b**) Comparison of the average transmission rate R of the system

average transmission rate of the OFDMA-VLC/RF system increases by 17 packages, compared with the VLC/RF system. This is because the OFDMA technology increases the capacity of the entire communication system, thereby increasing the system transmission rate.

Conclusion

In this study, the resource-allocation problem in an indoor OFDMA-VLC/RF system was investigated, and an allocation strategy based on the Lyapunov optimization algorithm was proposed. The purpose of this study was to solve the resource-allocation problem by maximizing the average transmission rate of the system under the constraints of system stability and an average power budget.

1. To find the optimal resource-allocation strategy, we formulated a random-optimization problem that considered the stability constraints and average power budget of the system, and used the average system transmission rate as the optimization goal.
2. Using the Lyapunov optimization algorithm, the stochastic optimization problem was decomposed into three independent subproblems, and their optimal solutions were provided. The effectiveness of the proposed resource-allocation strategies were verified through simulation.

The results showed that this strategy significantly improved the average transmission rate of the system, while maintaining the system stability. This means that our approach can improve the system performance by making more efficient use of limited resources in indoor OFDMA-VLC/RF systems. This study has practical significance for improving the performance of indoor wireless communication systems and provides a valuable reference for further research.

References

1. Y.Q. Zhou, L. Liu, L. Wang, et al. Service-aware 6G: an intelligent and open network based on the convergence of communication, computing and caching [J]. Digital Communications and Networks. 2020, 6 (3): 253–260.
2. D.Q. Ding, X.Z. Ke. Research on Visible Light Communication and Its Key Technologie [J]. Semiconductor optoelectronic. 2006(02):114–117.
3. M.A. Arfaoui, M.D. Soltani, I. Tavakkolnia, et al. Physical layer security for visible light communication systems: a survey [J]. IEEE Communications Surveys & Tutorials. 2020, 22 (3):1887–1908.
4. Wu W, Zhou F, Yang Q, et al. Adaptive Network Resource Optimization for Heterogeneous VLC/RF Wireless Networks [J]. IEEE Transactions on Communications, 2018, 66(11): 5568–5581.
5. V. K. Papanikolaou, P. D. Diamantoulakis, P. C. Sofotasios, et al. On optimal resource allocation for hybrid VLC/RF networks with common backhaul [J] IEEE Transactions on Cognitive Communications and Networking, 2020, 6(1): 352–365.

6. A. M. Abdelhady, O. Amin, A. Chaaban, et al. Downlink resource allocation for dynamic TDMA-based VLC systems [J]. IEEE Transactions on Wireless Communications, 2019, 18(1): 108–120.
7. M. Kashef, M. Abdallah, N. Al-Dhahir. Transmit power optimization for a hybrid PLC/VLC/RF communication system [J]. TGCN, 2018, 2(1): 234–245.
8. S. Aboagye, A. Ibrahim, T. M. N. Ngatched, et al. Design of Energy Efficient Hybrid VLC/RF/PLC Communication System for Indoor Networks [J]. IEEE Wireless Communications Letters, 2020, 9(2): 143–147.
9. Ma S, Zhang F, Li H, et al. Aggregated VLC-RF Systems: Achievable Rates, Optimal Power Allocation, and Energy Efficiency [J]. IEEE Transactions on Wireless Communications, 2020, 19(11): 7265–7278.
10. Obeed M, Salhab A M, Zummo S A, et al. Joint optimization of power allocation and load balancing for hybrid VLC/RF networks [J]. Journal of Optical Communications and Networking, 2018, 10(5): 553–562.
11. Pratama Y S M, Choi K W. Bandwidth Aggregation Protocol and Throughput-Optimal Scheduler for Hybrid RF and Visible Light Communication Systems [J]. IEEE Access, 2018, 6: 32173–32187.
12. Kahn J M, Barry J R. Wireless infrared communications [J]. Proceedings of the IEEE, 1997, 85(2): 265–298.
13. Carruther J B, Kahn J M. Angle diversity for nondirected wireless infrared communication [J]. IEEE Transactions on Communications, 2000, 48(6): 960–969.
14. Pang G, Ka-Lim H, Kwan T, et al. Visible light communication for audio systems [J]. IEEE Transactions on Consumer Electronics, 1999, 45(4): 1112–1118.
15. Pang G, Kwan T, Liu H, et al. Optical wireless based on high brightness visible LEDs [C]. Conference Record of the 1999 IEEE Industry Applications Conference. Thy-Forth IAS Annual Meeting (Cat. No.99CH36370), Phoenix, AZ, USA, 1999.
16. Tanaka Y, Haruyama S, Nakagawa M. Wireless optical transmissions with white colored LED for wireless home links [C]. 11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications. PIMRC 2000. Proceedings (Cat. No.00TH8525), London, UK, 2000.
17. Tanaka Y, Komine T, Haruyama S, et al. Indoor Visible Light Data Transmission System Utilizing White LED Lights [J]. IEICE Transactions on Communications, 2003, E86-B (8): p. 2440–2454.
18. Fan K, Komine T, Tanaka Y, et al. The effect of reflection on indoor visible-light communication system utilizing white LEDs [C]. The 5th International Symposium on Wireless Personal Multimedia Communications, Honolulu, HI, USA, 2002.
19. X.J. Che. Multipath characteristics analysis of indoor MIMO visible light communication channe MIMO [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2016.
20. Z.P. Sun. Research on network selection method and multi-user access technology in VLC-based indoor heterogeneous networks [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2018.
21. Hidemitsu, Sugiyama, Shinichiro, et al. Experimental Investigation of Modulation Method for Visible-Light Communications [J]. Ieice Transactions on Communications, 2006, 89-B (12): 3393–3400.
22. Afgani M Z, Haas H, Elgala H, et al. Visible light communication using OFDM [C]. 2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, 2006. TRIDENTCOM 2006., Barcelona, Spain, 2006.
23. Hashemi S K, Ghassemlooy Z, Chao L, et al. Orthogonal frequency division multiplexing for indoor optical wireless communications using visible light LEDs [C]. 2008 6th International Symposium on Communication Systems, Networks and Digital Signal Processing, Graz, Austria, 2008.
24. Langer K, Vucic J, Kottke C, et al. Advances and prospects in high-speed information broadcast using phosphorescent white-light LEDs [C]. 2009 11th International Conference on Transparent Optical Networks, Ponta Delgada, Portugal, 2009.

25. Vucic J, Kottke C, Nerreter S, et al. White Light Wireless Transmission at 200Mb/s Net Data Rate by Use of Discrete-Multitone Modulation [J]. IEEE Photonics Technology Letters, 2009, 21(20): 1511–1513.
26. Vučič J, Kottke C, Nerreter S, et al. 513 Mbit/s Visible Light Communications Link Based on DMT-Modulation of a White LED [J]. Journal of Lightwave Technology, 2010, 28(24): 3512–3518.
27. Kottke C, Hilt J, Habel K, et al. 803 Mbit/s visible light WDM link based on DMT modulation of a single RGB LED luminary [C]. 2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, Los Angeles, CA, USA, 2011.
28. Azhar A H, Tran T, Brien D O. Demonstration of high-speed data transmission using MIMO-OFDM visible light communications [C]. 2010 IEEE Globecom Workshops, Miami, FL, USA, 2010.
29. Zhang S, Watson S, Mckendry J J D, et al. 1.5 Gbit/s Multi-Channel Visible Light Communications Using CMOS-Controlled GaN-Based LEDs [J]. Journal of Lightwave Technology, 2013, 31(8): 1211–1216.
30. Tsonev D, Chun H, Rajbhandari S, et al. A 3-Gb/s Single-LED OFDM-Based Wireless VLC Link Using a Gallium Nitride [J]. IEEE Photonics Technology Letters, 2014, 26(7): 637–640.
31. Lu I, Yeh C, Hsu D, et al. Utilization of 1-GHz VCSEL for 11.1-Gbps OFDM VLC Wireless Communication [J]. IEEE Photonics Journal, 2016, 8(3): 1–6.
32. Yeh C H, Weng J H, Chow C W, et al. 1.7 to 2.3 Gbps OOK LED VLC Transmission Based on 4×4 Color-Polarization-Multiplexing at Extremely Low Illumination [J]. IEEE Photonics Journal, 2019, 11(4): 1–6.
33. Gunawan W H, Liu Y, Yeh C H, et al. Color-Shift-Keying Embedded Direct-Current Optical-Orthogonal-Frequency-Division-Multiplexing (CSK-DCO-OFDM) for Visible Light Communications (VLC) [J]. IEEE Photonics Journal, 2020, 12(5): 1–5.
34. Lin B, Lai Q, Ghassemlooy Z, et al. A Machine Learning Based Signal Demodulator in NOMA-VLC [J]. Journal of Lightwave Technology, 2021, 99, 1-1.
35. Abdelhady A M, Amin O, Chaaban A, et al. Downlink Resource Allocation for Dynamic TDMA-Based VLC Systems [J]. IEEE Transactions on Wireless Communications, 2019, 18(1): 108–20.
36. Aboagye S, Ngatched T M N, Dobre O A, et al. Energy Efficient Subchannel and Power Allocation in Cooperative VLC Systems [J]. IEEE Communications Letters, 2021, 99, 1-1.
37. Fu Y, Hong Y, Chen L, et al. Enhanced Power Allocation for Sum Rate Maximization in OFDM-NOMA VLC Systems [J]. IEEE Photonics Technology Letters, 2018, 30(13): 1218–1221.
38. Li X, Zhang R, Hanzo L. Cooperative Load Balancing in Hybrid Visible Light Communications and WiFi [J]. IEEE Transactions on Communications, 2015, 63(4): 1319–1329.
39. Mohanad O, Salhab A M, Zummo S A, et al. Joint Optimization of Power Allocation and Load Balancing for Hybrid VLC/RF Networks [J]. IEEE/OSA Journal of Optical Communications & Networking, 2018, 10(5): 553–662.
40. Kafafy M, Fahmy Y, Abdallah M, et al. A novel bandwidth and power allocation scheme for power efficient hybrid RF/VLC indoor systems [J]. Physical Communication, 2018, 31(DEC.): 187–195.
41. Kashef M, Ismail M, Abdallah M, et al. Energy Efficient Resource Allocation for Mixed RF/VLC Heterogeneous Wireless Networks [J]. IEEE Journal on Selected Areas in Communications, 2016, 34(4): 883–893.
42. Dakhli C M, Zayani R, Belkacem B O, et al. Theoretical analysis and compensation for the joint effects of HPA nonlinearity and RF crosstalk in VBLAST MIMO-OFDM systems over Rayleigh fading channel [J]. EURASIP Journal on Wireless Communications and Networking, 2014, 2014(1): 1–15.
43. Hammouda M, Vegni A M, Haas H, et al. Resource Allocation and Interference Management in OFDMA-based VLC Networks [J]. Physical Communication, 2018, 31(Dec.):169–180.

44. Armstrong J, Schmidt B J C. Comparison of Asymmetrically Clipped Optical OFDM and DC-Biased Optical OFDM in AWGN [J]. IEEE Communications Letters, 2008, 12(5): 343–345.
45. Wang Y, He Y, Xu C, et al. Joint rate control and power allocation for low-latency reliable D2D-based relay network [J]. EURASIP Journal on Wireless Communications and Networking, 2019, 2019(1):1–14.
46. W. Bao, H. Chen, Y. Li, et al. Joint rate control and power allocation for non-orthogonal multiple access systems [J]. IEEE Journal on Selected Areas in Communications, 2017, 35(12): 2798–2811.
47. Y. Guo, Q. Yang, K. S. Kwak, et al. Quality-oriented rate control and resource allocation in time-varying OFDMA networks [J]. IEEE Transactions on Vehicular Technology, 2017, 66(3): 2324–2338.
48. R. Zhang, H. Claussen, H. Haas, et al. Energy efficient visible light communications relying on amorphous cells [J]. IEEE Journal on Selected Areas in Communications, 2016, 34(4):894–906.
49. Valencia-estrada J C, García-marquez J, Chassagne L, et al. Catadioptric interfaces for designing VLC antennae [J]. Applied Optics, 2017, 56(27): 7559.
50. S. A. AlQahtani, M. Alhassany. Comparing Different LTE Scheduling Schemes [C]. IEEE 9th International Wireless Communications and Mobile Computing Conference, 2013, 264–269.

Indoor Visible-Light Location Based on a Fusion Clustering Algorithm

41

Contents

Introduction 1528
Indoor Model and Algorithm Principle 1530
System Model 1530
Channel Model 1530
Fingerprint-Location Method 1532
Improved Fingerprint Method 1533
Offline Phase 1533
Online Phase 1535
Simulation Verification 1536
Fingerprint Method 1537
Improved Fingerprint Method 1539
Discussion 1544
Conclusions 1545
References 1546

Abstract

A new fingerprint-location method based on a clustering algorithm was proposed to address the problem of large positioning errors in an indoor boundary area. This clustering-based method embodies the idea of a rough location and then a fine location. First, the edge regions of the received signal strength (RSS) samples, which are greatly affected by reflection, are divided using the k-medoids algorithm. Then, the center part is clustered using density-based spatial clustering of applications with noise (DBSCAN). In the actual location-estimation stage, the points to be measured can only be located in one of the classified areas and are combined with the weighted-optimum k-nearest-neighbor algorithm (WOKNN) to match the location. The results showed that the average positioning error of the algorithm was 13 cm in an indoor environment measuring 5 m × 5 m × 3 m. Compared with the traditional method without clustering, the positioning accuracy of the edge area increased by 21%, and the overall

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_41

improvement was 33.8%. This proves that the proposed algorithm effectively improves the efficiency of real-time indoor-positioning accuracy.

Keywords

Indoor positioning · Visible-light positioning · Region division · Position fingerprint method

Introduction

With increasing improvements in the Internet of Things, the application of location-based services has attracted considerable attention. Large-scale indoor construction is increasing gradually, and the demand for indoor positioning is increasing rapidly. In recent years, indoor positioning has been widely used in urban construction, medical services, construction sites, chemical metallurgy, power plants, and other fields. In the architecture field, the main applications include emergency management [1], intelligent energy management [2], heating, ventilation, air-conditioning system (HVAC) controls [3], and occupancy detection [4].

Real-time accurate knowledge of the locations of people in a building is useful for a variety of applications. At present, indoor wireless-positioning technologies are being developed, including wireless local-area networks (WLANs), radio frequency identification (RFID), Bluetooth, ultrawideband (UWB), ultrasonic, and infrared. Among these technologies, WiFi-based and Bluetooth-based low-energy (BLE) location technologies have become the most commonly used and are widely deployed in current smart devices, mainly because they do not require a line-of-sight path between the transmitter and receiver [5].

However, WiFi location relies on the assumption that the signal strength at each point in space does not change with time or other obstacles, which is unrealistic in real-world scenarios, and WiFi-location results are often unreliable. Multiple-input multiple-output (MIMO) communication has long played an important role in RF-based communication. Using MIMO, data rates can be increased without increasing the bandwidth or power, thereby increasing the channel capacity and effectively supporting multiple users [6].

In contrast to the aforementioned indoor-positioning technologies based on radio frequency communication, visible-light positioning (VLP) technology is based on indoor visible-light communication (VLC) technology. Indoor positioning can be realized using light-emitting diode (LED) lamps for lighting. VLC has developed rapidly, owing to its advantage of being green, its resistance to electromagnetic interference, and its rich spectrum resources. It can be used to enhance future wireless communication systems and has great potential for providing reliable solutions to indoor-positioning problems.

Sejan studied VLC technology in a MIMO communication environment and considered a machine learning method for MIMO VLC [7]. To construct a VLP system, methods such as the angle of arrival (AOA), time of arrival (TOA), and

received signal strength (RSS) can be used. TOA requires the strict synchronization of transmission and reception, whereas AOA requires more complex hardware. Therefore, methods based on RSS have attracted extensive attention, owing to its low cost and simple hardware [8].

Previous research based on VLC ignored the reflections of ceilings, walls, and indoor objects, making the positioning effect too ideal and inconsistent with results in an actual environment. In 2016, Aminikashani demonstrated the influence of multipath reflections and found that the central area of a room was almost unaffected. In contrast, other areas (viz., corners and edges) had very poor effects [9]. Saadi M observed that using a clustering algorithm before the location process meant that the process of querying the closest point only occurred in one area, which shortened the total execution time and improved the location accuracy [10].

In 2016, Saadi used only two LEDs and the k-means clustering method and obtained an average localization error of 31 cm, which was the best accuracy achieved in any practical environment using LEDs [11]. At present, k-means is widely used in indoor positioning; however, its initial centroid selection is random, which causes the clustering to find a local optimal solution, resulting in poor clustering results.

In 2018, Wang studied a visible-light positioning system model based on the binary k-means clustering method, which improved the clustering effect and simplified the computational complexity [12]. Dividing by region guaranteed the signal integrity in each region; however, the fingerprint points in the edge region of the cluster could be divided into incorrect clusters. In 2022, Ren proposed a hybrid localization algorithm based on an extreme learning machine (ELM) and density-based spatial clustering of applications with noise (DBSCAN) to solve the problem of large boundary errors between edges and corner clusters. Through comparison, it was found that the DBSCAN clustering algorithm was the best maximum-likelihood localization algorithm [13].

In the same year, Martínez proposed a new VLP system with RGB LEDs for multicell networks. It used the Euclidean distance and RGB LED position to estimate the position of the receiver through a trilateration algorithm. This system could be extended to multiple units based on the application scenario and had practical applications [14]. In 2023, to better realize real-time positioning, Long divided the positioning points into ordinary points, edge points, and blind spots and used corresponding models to handle the different types of points. This was the first study to consider the different effects of multipath effects in different areas of a room and realized centimeter-level positioning [15].

Improving the accuracy of indoor visible-light positioning systems using a simple, real-time, and robust approach has been a challenge in recent research [16]. In this study, we consider a VLC channel model that includes line-of-sight (LOS) and nonline-of-sight (NLOS) components. Considering that the randomness of k-means clustering greatly destabilizes the location, a fusion method combining k-medoids clustering with DBSCAN is proposed. It has low algorithm complexity and can ensure that fingerprints in the same environment, at different times and at the same location, have high similarity. Then, the fusion clustering algorithm was

compared with the weighted-optimum k-nearest-neighbor (WOKNN) algorithm to discuss the localization of different regions and verify the robustness and effectiveness of the algorithm.

Indoor Model and Algorithm Principle

System Model

An indoor visible-light positioning system based on RSS is mainly composed of an LED light source, a visible-light channel, and a photodetector (PD) array, as shown in Fig. 41.1. In this study, the room size was set to 5 m × 5 m × 3 m, and four LED arrays were evenly and symmetrically installed on the ceiling. They were used as emission sources and located at (1, 1, 3), (1, 4, 3), (4, 1, 3), and (4, 4, 3). The receiving end faced horizontally upward from any position on the ground.

Channel Model

In this study, considering the actual indoor environment, a line-of-sight (LOS) link and a nonline-of-sight (NLOS) link with primary wall reflection were considered. The communication link is shown in Fig. 41.2.

Assuming that the luminous intensity of the LED follows the Lambert radiation model, the channel gain in the LOS link can be expressed as [17]

$$H_{\mathrm{LOS}}(0) = \begin{cases} \dfrac{(m+1)A_r \cos^m(\phi)\cos(\varphi)}{2\pi \mathrm{d}^2} \mathrm{T}(\varphi)\mathrm{g}(\varphi) & 0 \le \varphi \le \varphi_{\mathrm{FOV}} \\ 0 & \varphi > \varphi_{\mathrm{FOV}} \end{cases}, \tag{41.1}$$

where A_r is the physical area of the photodetector (PD) and m is the Lambert radiation order. d is the distance between the LED and PD, φ is the incident angle,

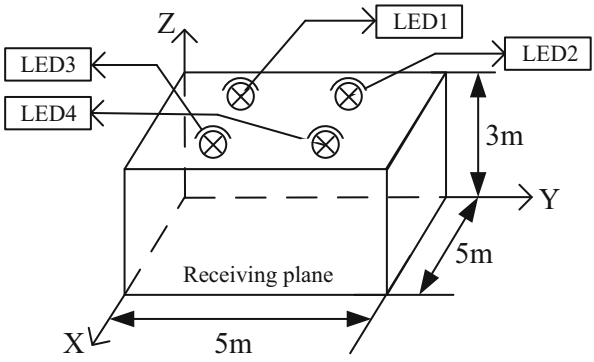


Fig. 41.1 Indoor VLC system model [16]

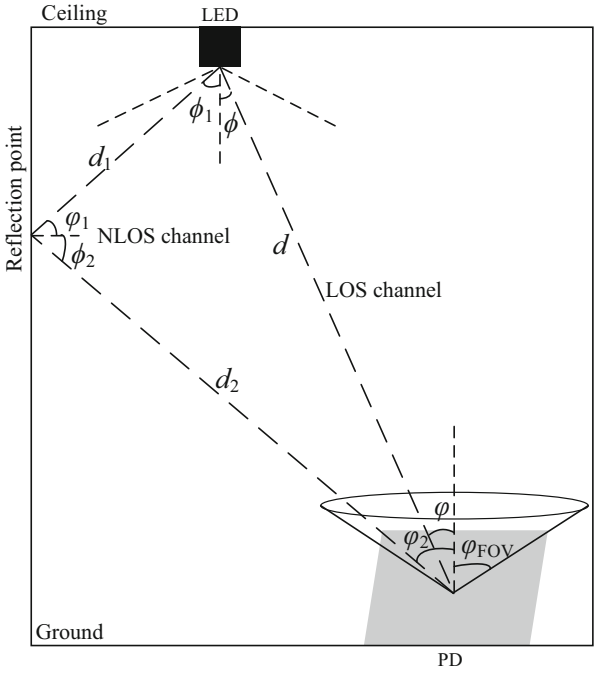


Fig. 41.2 Indoor visible-light channel [18]

ϕ is the irradiance angle, and φ_{FOV} is the field of view of the receiver. $T(\varphi)$ and $g(\varphi)$, respectively, represent the gains of the optical filter and optical condenser of the receiver, which can be expressed as

$$g(\varphi) = \begin{cases} \dfrac{n^2}{\sin^2(\varphi_{\mathrm{FOV}})} & 0 \le \varphi \le \varphi_{\mathrm{FOV}} \\ 0 & \varphi > \varphi_{\mathrm{FOV}} \end{cases}, \tag{41.2}$$

where n is the refractive index of the optical concentrator.

Assuming that the wall consists of many diffuse mirrors with area $\triangle$A and reflectivity $\mathrm{R_W}$, the channel gain of the first-order reflection is given by the following formula:

$$H_{\mathrm{dif}}(0) = \begin{cases} \sum\limits_{\mathrm{Wall}} \dfrac{A_{\mathrm{r}} \Delta A R_{\mathrm{W}}(m+1)}{2\pi^2 d_1^2 d_2^2} \cos^m(\phi_1)\cos(\varphi_1)\cos(\phi_2)\cos(\varphi_2)T(\varphi_2)g(\varphi_2) & 0 \le \varphi_2 \le \varphi_{\mathrm{FOV}} \\ 0 & \varphi_2 > \varphi_{\mathrm{FOV}} \end{cases}, \tag{41.3}$$

where d_1 and d_2 are distances, φ_1 and φ_2 are incident angles, ϕ_1 and ϕ_2 are irradiance angles, R_W is the reflectivity of the wall, and $\triangle$A is the area element of the reflection point of the wall.

The final total received power is expressed as

$$P_{\mathrm{rs}} = \{RP_{\mathrm{t}}[H_{\mathrm{LOS}}(0) + H_{\mathrm{dif}}(0)]\}^2 + \sigma^2_{\mathrm{noise}}, \tag{41.4}$$

where P_{rs} is the received power, P_t is the transmitted power, R is the responsivity of the receiver, and $\sigma^2_{\mathrm{noise}}$ is the variance of the total noise.

In a typical VLP system, LEDs are modulated to transmit signals, and time division multiplexing (TDM), frequency division multiplexing (FDM), or code division multiplexing (CDM) is used to allow multiple accesses from different LEDs. Therefore, intersymbol interference (ISI) can be avoided and background noise can be filtered out. The variance of noise can be expressed as [19]

$$\begin{aligned}\sigma^2_{\mathrm{noise}} &= \sigma^2_{\mathrm{shot}} + \sigma^2_{\mathrm{thermal}} \\ &= (2qRP_{rs}B) + \left(\frac{8\pi kT_k}{G}\eta A_r I_2 B^2 + \frac{16\pi^2 kT_k \Gamma}{g_m}\eta^2 A_r^2 I_3 B^3\right),\end{aligned} \tag{41.5}$$

where q, B, and k represent the charge, equivalent noise bandwidth, and Boltzmann constant, respectively, $k = 1.38 \times 10^{-23}$. T_k, G, and η represent the absolute temperature, open loop voltage gain, and fixed PD capacitance, respectively. Γ and g_m represent the field-effect transistor (FET) channel noise figure and FET transconductance, respectively. I_2 and I_3 are 0.562 and 0.0868, respectively.

The signal-to-noise ratio is expressed as [20]

$$SNR = \frac{P_r^2}{\sigma^2_{\mathrm{noise}}} = \frac{\{RP_t[H_{\mathrm{LOS}}(0) + H_{\mathrm{dif}}(0)]\}^2}{\sigma^2_{\mathrm{shot}} + \sigma^2_{\mathrm{thermal}}}, \tag{41.6}$$

where P_r^2 is the signal power of the receiver.

Fingerprint-Location Method

The fingerprint-location method infers the final location of the target through the physical-location coordinates and its signal characteristics (such as the RSS value). It uses a matching algorithm, which better inhibits the multipath effect and nonline-of-sight signal propagation [21]. It includes offline (training) and online (testing) modes, as shown in Fig. 41.3.

The main task of the offline mode is to collect the RSS from all fingerprints; each physical location should have a unique and distinguishable fingerprint. Then, noise reduction, regional division, and training are carried out in turn, and a fingerprint database is established according to certain rules. The online mode uses the training data in the offline stage to predict the current position of the PD. Existing matching algorithms include the nearest neighbor (NN), k-nearest neighbor (KNN), weighted k-nearest neighbor (WKNN), neural network, and support vector machine (SVM) [23].

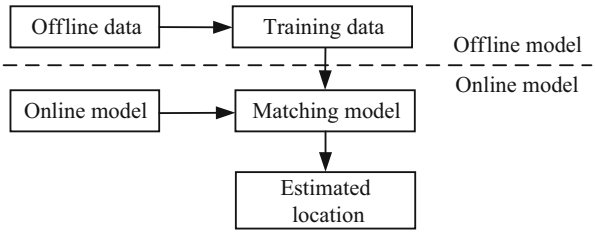


Fig. 41.3 Indoor-positioning system based on a fingerprint [22]

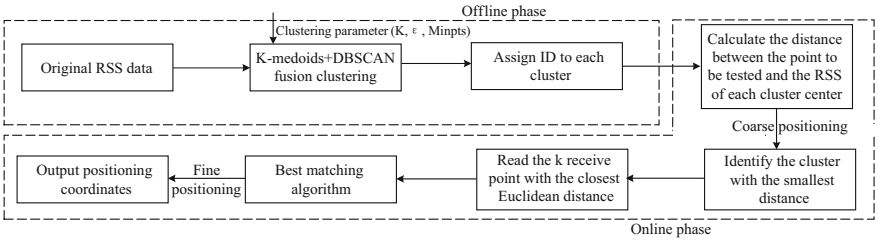


Fig. 41.4 Block diagram of the proposed algorithm

Improved Fingerprint Method

Research on fingerprint technology can be divided into two topics: establishing a fingerprint database in the offline stage and optimizing the matching algorithm in the online stage. A block diagram of the proposed algorithm is shown in Fig. 41.4.

Considering the comprehensive characteristics of the RSS in the offline stage, a fusion clustering algorithm was proposed. In the online-positioning stage, the Euclidean distance between the point to be measured and the RSS value of the central node of each cluster was calculated. The positioning area was narrowed to the area where the cluster closest to the central node was located, and the coordinate information of the positioning point was estimated using the best matching model to complete the positioning.

Offline Phase

The purpose of offline clustering is to reduce the space and computational complexity of a fingerprint search. The most important aspect is to improve the correctness of the fingerprint matching and location accuracy [24]. The proposed fusion clustering is as follows. The k-medoids algorithm divides the edge and middle regions, and the DBSCAN algorithm distributes the middle part according to the light intensity distribution characteristics. Feature extraction is important in NLOS recognition and mitigation.

The clustering effect is often measured using the Davies–Bouldin (DB) index, silhouette index (SI), and the sum of squares error (SSE). The smaller the DB or SSE, or the larger the SI, the closer the data objects are to their centroids, and the better the clustering effect [25].

K-medoids Algorithm

The k-medoids algorithm was optimized using k-means clustering. K-means takes the center point as the average of all data points in the current cluster, whereas k-medoids selects the point with the smallest distance to all other points in the current cluster as the central point. This solves the problems of sensitivity to abnormal points and being greatly affected by extreme values.

The algorithm flowchart is shown in Fig. 41.5, and the detailed steps are as follows [26]:

Randomly select k points from the total n sample points as medoids. Assign each of the remaining n-k points to the class of the closest medoid. For each point in class i, except the medoid point, calculate the value of the cost function. Repeat this with the other classes. Then, for each class, select the point with the minimum cost function as the new medoid. Repeat the first two processes until the medoid points no longer change or the set maximum number of iterations has been reached. Then, output the final k classes.

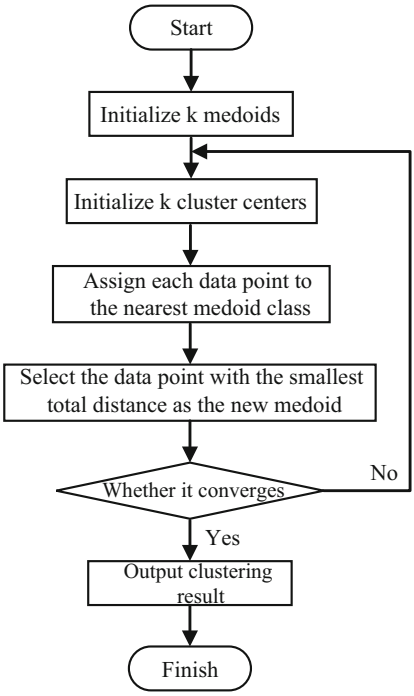


Fig. 41.5 K-medoids clustering algorithm flowchart

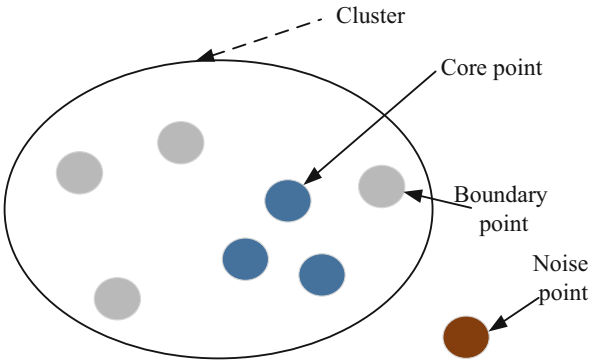


Fig. 41.6 Examples of core points, boundary points, and noise points [22]

DBSCAN Algorithm

DBSCAN, proposed by Ester in 1996, is a nonparametric density-based clustering algorithm. Its advantage is that it does not need to specify the number of clusters in advance. It divides high-density points into clusters and can effectively deal with low-density noise points [27]. Using the DBSCAN algorithm, the points of the fingerprint dataset were divided into core, boundary, and noise points. As shown in Fig. 41.6, the type to which each point belongs depends on the sample distribution, *Eps*, and *Minpts*. *Eps* is the radius of the region used to determine whether the data points are located in the cluster and *Minpts* specifies the minimum number of points required to form the cluster. The initial selection of *Eps* and *Minpts* greatly influences the clustering effect [28].

The DBSCAN process is as follows [29]: Starting from any point p in the database, retrieve all point densities reachable from p under the given *Eps* and *Minpts*; the point density is the compactness of the sample point distribution. If p is the core point, the process generates a cluster using *Eps* and *Minpts*, and the point is classified. If p is a boundary point, there is no point density reachable from p, and DBSCAN accesses the next unclassified point in the database [30]. Because DBSCAN defines clusters based on density, it is relatively noise resistant and can find clusters of arbitrary shapes. Therefore, DBSCAN can identify clusters that k-means clustering cannot.

Online Phase

The offline clustering results are stored in a new location fingerprint database. In the positioning stage, the distance between the label to be tested and the central node of each subregion is obtained for rough positioning. Subsequently, fine positioning is completed by matching the coarse positioning area. Current matching algorithms include NN, KNN, WKNN, and simple weighted k-nearest neighbor (SWKNN) [31]. To maximize its advantages, this study implemented a weighted optimal

k-nearest neighbor (WOKNN) algorithm in the positioning stage. This is the method for selecting the best matching algorithm and its best k value for different regions.

The flow of the SWKNN algorithm is as follows:

1. The receiver obtains the RSS value and calculates its Euclidean distance d_i:

$$d_i = \sqrt{\sum_{i=1}^{4} (\mathrm{RSS}_{Ti} - \mathrm{RSS}_i)^2}, \tag{41.7}$$

where RSS_{Ti} represents the RSS value collected online at an undetermined site, RSS_i is the corresponding number from the offline fingerprint database, and i is the subscript of the LED.

2. The distances d_i are arranged in ascending order and the average value E_d of the nearest K distances are determined:

$$E_d = \frac{d_1 + d_2 + \cdots + d_k}{K}. \tag{41.8}$$

3. Each distance value d_i is compared with the average value E_d. Distances greater than the average value are removed, the remaining points are recorded as M, and the K value in WKNN is replaced with M.
4. The above process is repeated. The value of K is gradually reduced and gradually approaches the true value.

The estimated position coordinates of the target are

$$P = \frac{\frac{1}{d_1}L_1 + \frac{1}{d_2}L_2 + \cdots + \frac{1}{d_M}L_M}{\frac{1}{d_1} + \frac{1}{d_2} + \cdots + \frac{1}{d_M}}, \tag{41.9}$$

where P represents the coordinates of the target to be located and L_i is the position of the i^{th} fingerprint.

Simulation Verification

This section describes the simulation results obtained using the proposed method. Starting from the power distribution and considering the influence of diffuse reflections, the positioning error distribution in different areas of an indoor environment was studied to verify the effectiveness of the proposed algorithm. The additional conditions are summarized in Table 41.1. Figure 41.7 shows the indoor illumination

intensity distribution on the same plane, based on the Lambertian model. It conforms to the international standard for indoor office illumination between 300 and 1500 lx.

Fingerprint Method

The LOS channel is very important in almost all indoor VLC positioning systems because it receives most of the total optical power. Figure 41.8a shows the power distribution under the LOS link; the entire receiving surface presents a high center and low corners.

Table 41.1 VLP system simulation parameters

Parameters	Values
Emitting optical power	1 W
Half power angle	60°
Wall reflectivity	0.7
Receiver responsiveness	0.5 A/W
Receiver field-of-view angle	70°
Refractive index	1.5
Reflection coefficient	0.8
Signal-to-noise ratio (SNR)	30 dB

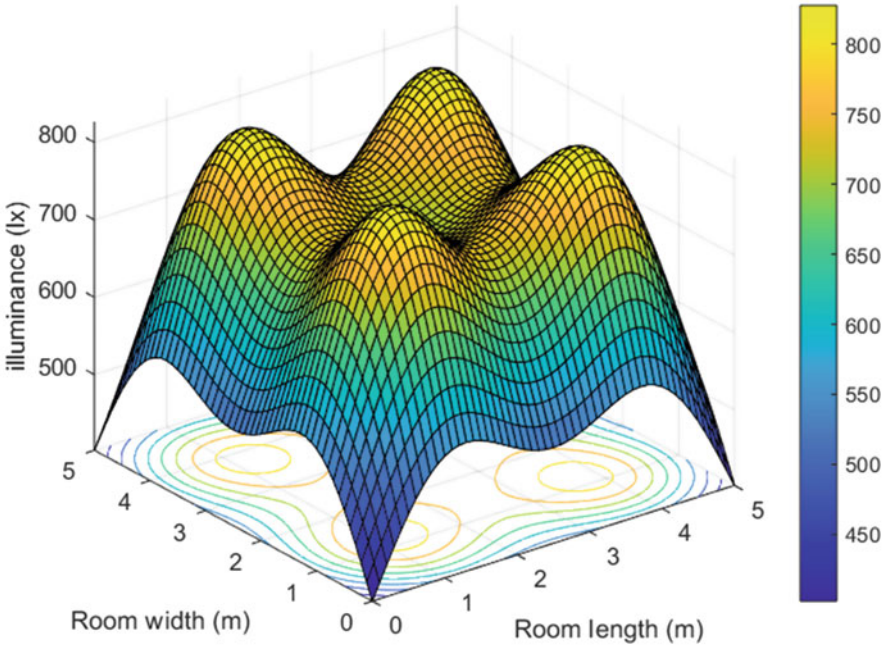


Fig. 41.7 Indoor communication receiving plane light intensity distribution

Figure 41.8b shows the optical-power distribution diagram of the NLOS. It can be seen that the distributions of the NLOS and LOS channels show opposite trends. Visible light always reflects in indoor environments, especially in narrow spaces with walls, ceilings, floors, and furniture. Therefore, most of the power added by the NLOS of the scattering part is concentrated near the walls of the room, with less in the center; the center receives the lowest scattered power. Compared with Fig. 41.8a, the scattering power near the edge of the wall is almost equal to that of the LOS component.

Figure 41.8c shows the distribution of the total received optical power. Compared with Fig. 41.8a, the overall power distribution of the room is wider, owing to the addition of NLOS energy.

When the SNR is 30 dB, the distribution of the positioning error of the 100 points to be measured is shown in Fig. 41.9. The numbers marked in the figure are the serial numbers of the points to be measured, which are evenly selected. It can be observed that the error in the middle area of the room is small, whereas the errors in the corner and edge areas are large. This is due to the increase in error caused by the gradual increase in the specific gravity of the scattering component; the multipath effect greatly increases the positioning error of the corners and edges.

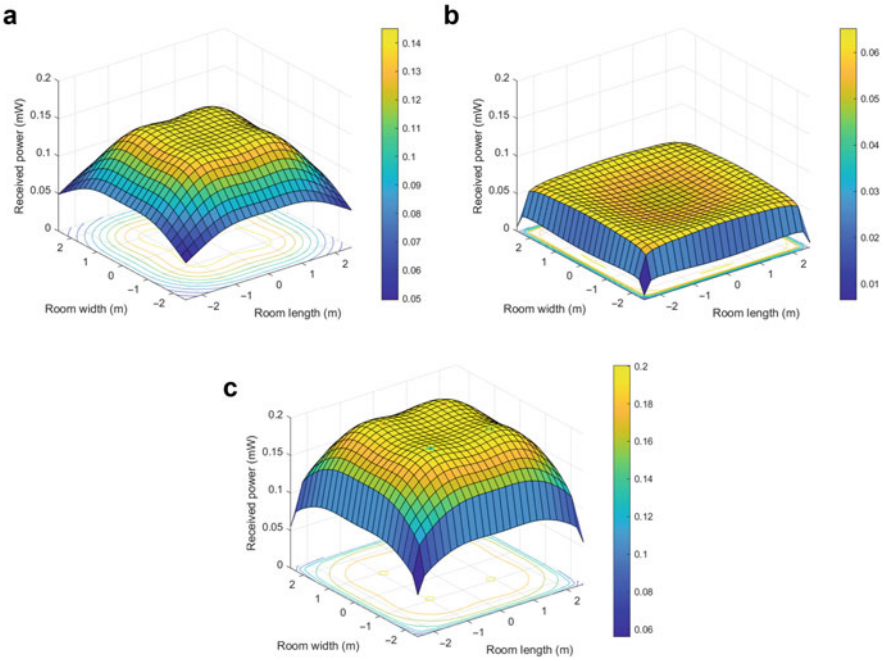


Fig. 41.8 Distribution of received optical power: (**a**) LOS channel, (**b**) channel with only one reflection, and (**c**) total channel

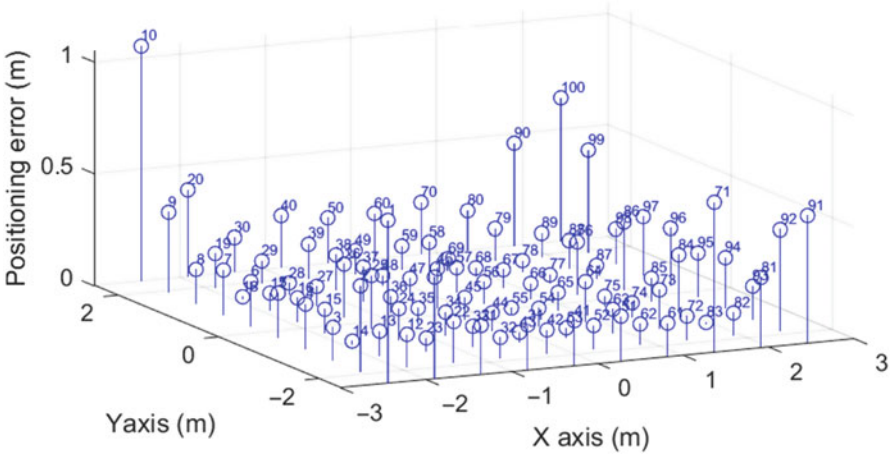


Fig. 41.9 Location error using the fingerprint method

According to the system simulation calculation, the average positioning error of the entire room was 20.26 cm, and the maximum error at the corners, which was greatly affected by the reflected light, was approximately 1.15 m. Thus, the traditional fingerprint positioning method was improved, and its performance was verified.

Improved Fingerprint Method

When correctly dividing clusters, the silhouette index (SI), elbow method, and sum of squares error (SSE) can be used as criteria for judging the clustering effect of k-medoids. They can be combined to determine the optimal number of clusters and initial K value. The initial selection of clustering centers is crucial to the clustering effect.

This study used the elbow method based on slope determination to determine the optimal number of clusters, k. The value of k in the K-medoids algorithm is three and the distance measure is a city block. In this case, the clustering effect was consistent with expectations and effectively distinguished between the center and edge areas.

For DBSCAN clustering, the clustering validity index was used to determine the optimal clustering result. The indices included the Davies–Bouldin (DB) index, silhouette index (SI), and sum of squares error (SSE). Among these, the SI index is widely used because of its simplicity, ease of use, and good evaluation capabilities. After simulation verification, the best clustering effect was determined. The final result was that when the number of clusters was five, the best *Minpts* value was 49 and the best *Eps* value was 1.5879×10^{-5}. These optimized parameters were used for DBSCAN clustering.

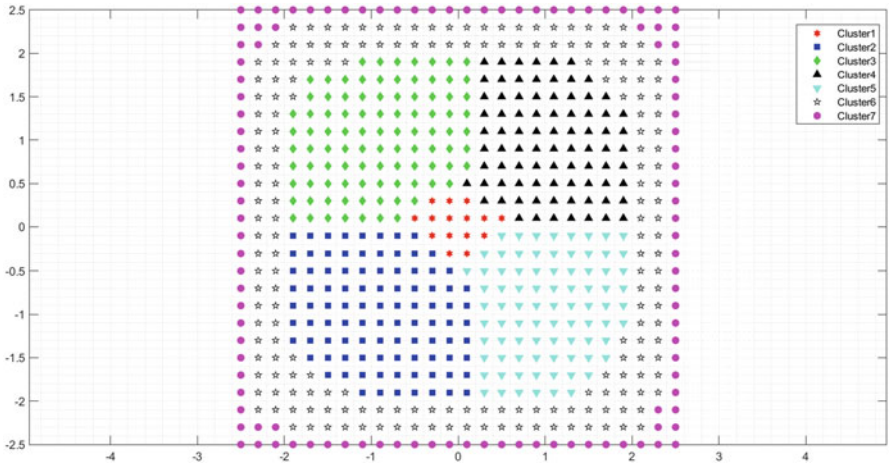


Fig. 41.10 Fusion clustering result diagram

The positioning effects differ in different areas. Therefore, VLC-based indoor-positioning technology must consider the reflected signals, and all receiving points are first positioned through fusion and clustering. Fusion clustering uses two algorithms to complement each other's advantages and improve the accuracy. The specific fusion clustering effect is shown in Fig. 41.10.

In the matching positioning stage, the value of k is manually selected and fixed [32]. The selected value may be well suited for one area, but other areas may require different values. Therefore, optimizing the number of nearest points in each region is a promising solution for improving positioning accuracy.

Subsequently, we simulated and analyzed the positioning error for different k values under different algorithms in each area, as shown in Fig. 41.10. For each cluster, from the perspective of overall positioning performance, SWKNN is the best, followed by WKNN and finally KNN. As the value of k increases, the error of the matching algorithm increases, which increases the positioning time; thus, the influence of the k value cannot be ignored. To minimize the average positioning error, it is necessary to determine the algorithm and its neighboring points with the minimum errors in each region.

Through cross-validation, the algorithm and k value with the best positioning effect in each cluster were determined. The results are summarized in Fig. 41.11 and presented in Table 41.2. From the center of the room to the edge, the average positioning error increases from small to large values, and the error in each area is reduced using the clustering algorithm.

Figure 41.12 shows the error distribution of the uniformly selected points to be measured. Compared with Fig. 41.7, it can be seen that the proposed algorithm effectively reduces the error of the edge part, with a maximum reduction of 0.49 m. It is more accurate for the edge region, partly because of clustering and partly because the cosine distance was used to measure the points to be measured instead of the

traditional Euclidean distance; this can make the estimated value closer to the real value.

The positioning accuracy is most obviously improved at the corners after adjustment. The positioning error in the central area of the room shows little improvement because it is not greatly affected by multipath reflection. This scheme can effectively solve the defects of large edge and corner errors in traditional RSS positioning, and the proposed algorithm greatly enhances the positioning performance based on VLP.

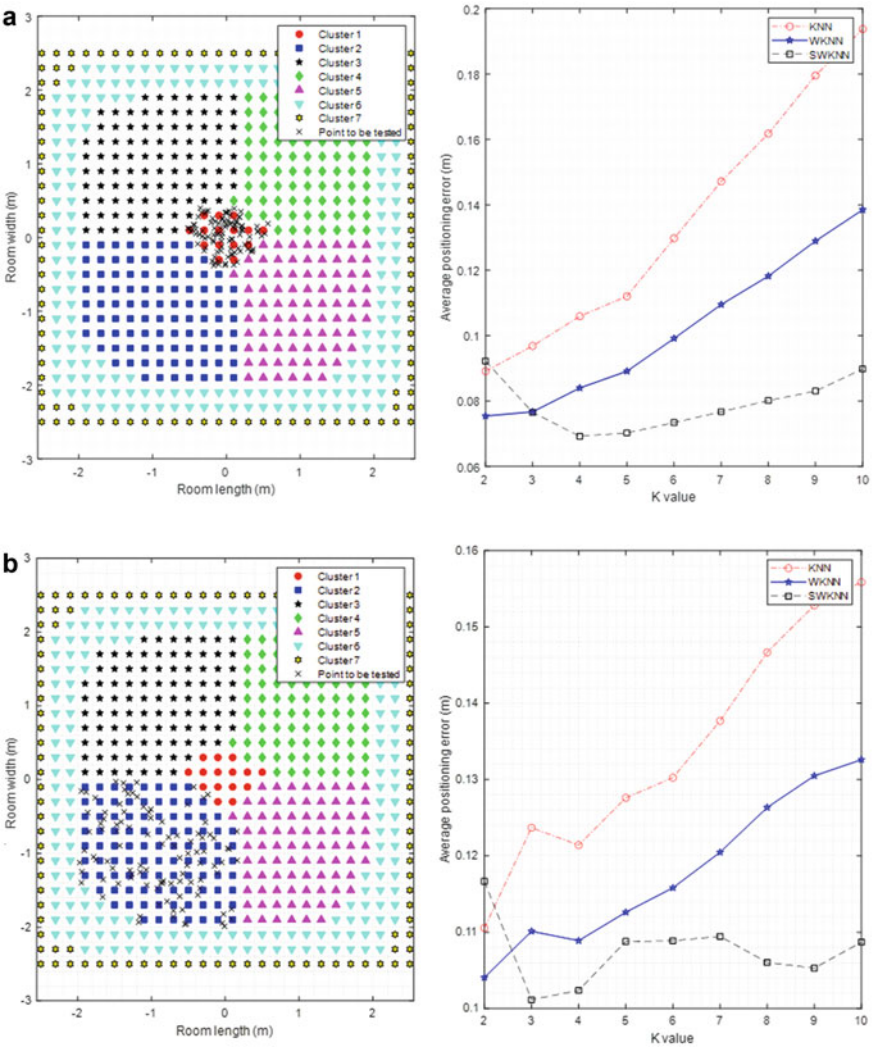


Fig. 41.11 Clustering effect diagram and positioning error of each region. (**a**) Cluster 1. (**b**) Cluster 2. (**c**) Cluster 3. (**d**) Cluster 4. (**e**) Cluster 5. (**f**) Cluster 6. (**g**) Cluster 7

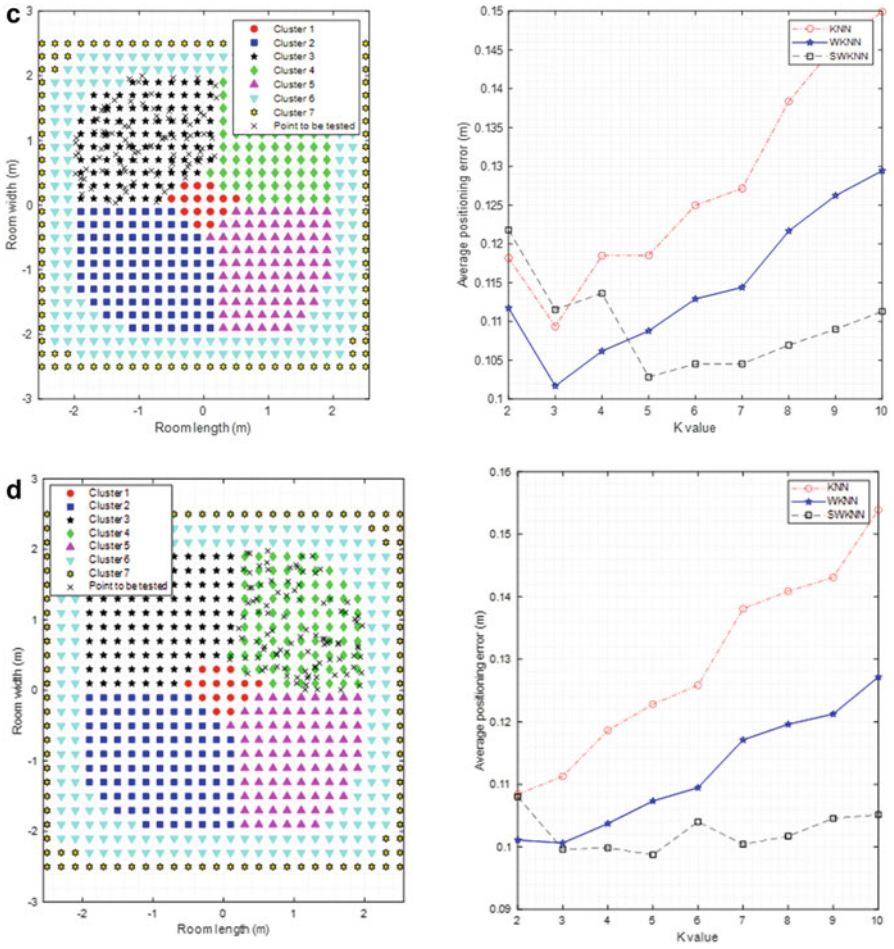


Fig. 41.11 (continued)

To further evaluate the localization performance of the proposed system, it was compared with classic algorithms. Figure 41.13 shows the cumulative distribution function (CDF) of the positioning errors for the different algorithms. The results show that the combination of the fusion clustering algorithm and WOKNN algorithm has a higher positioning accuracy than the other algorithms. When the cumulative positioning error distribution is 90%, the average positioning error of NN is 0.43 m, while the error distance of the proposed algorithm is 0.24 m. The positioning performance is improved by 44.2%, which is obviously improved as a whole.

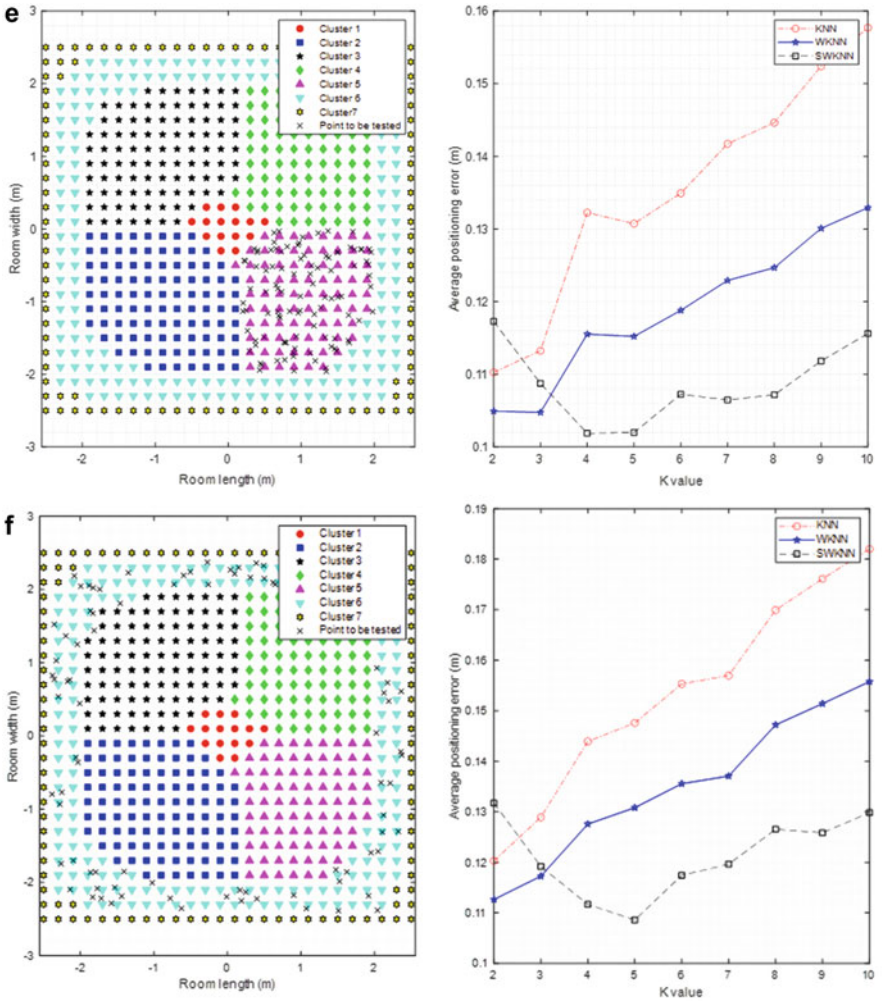


Fig. 41.11 (continued)

Table 41.3 indicates that the average positioning effect of the proposed algorithm is better than those of the other four more traditional algorithms. Using the proposed positioning algorithm, the positioning error at the training point can be effectively reduced, and the positioning accuracy of the entire receiving surface can be improved by 33.8%. This demonstrates the effectiveness of the proposed method. The improved algorithm proposed in this study greatly improved the clustering accuracy.

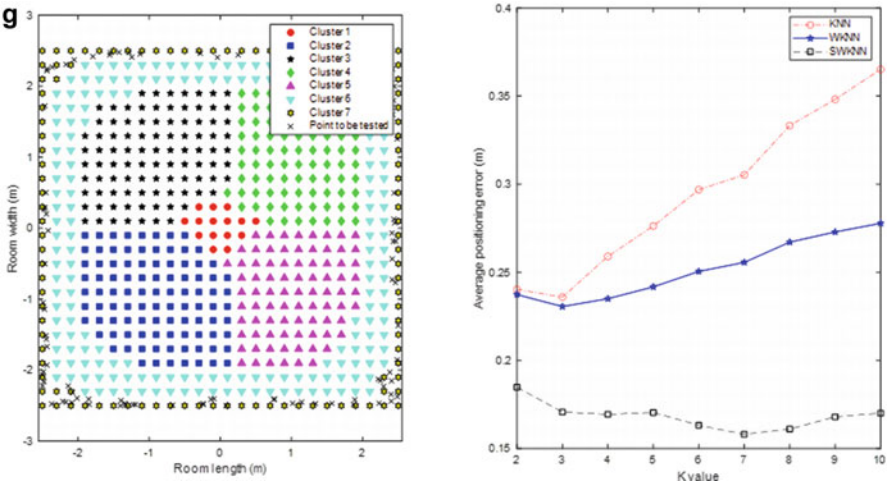


Fig. 41.11 (continued)

Table 41.2 Location results of different clusters

Region	Optimal algorithm	Optimum *k* value	Average positioning error (cm)
Cluster 1	SWKNN	4	6.92
Cluster 2	SWKNN	3	10.11
Cluster 3	WKNN	3	10.16
Cluster 4	SWKNN	5	9.87
Cluster 5	SWKNN	4	10.18
Cluster 6	SWKNN	5	10.86
Cluster 7	SWKNN	7	15.80

Discussion

The improved fingerprint method divides the research area into multiple clusters and analyzes each cluster separately to determine the optimal algorithm and k value. The algorithm proposed in this study is effective for most light source distributions with central symmetry. However, as indoor environments become more complex, the changing environment will definitely increase the failure of the fingerprint database. The focus should be on improving the maintenance and updating the efficiency of the fingerprint database. By selecting a special location as a monitoring point for detecting fingerprint data fluctuations, a backup fingerprint library for common environments can be developed to update the fingerprints over time.

Clustering is also conducive to the analysis of local features to a certain extent and can provide a reference for future compensation schemes. Therefore, to improve the practicability of indoor visible-light positioning, fingerprint database maintenance and updating technologies should be strengthened in the future.

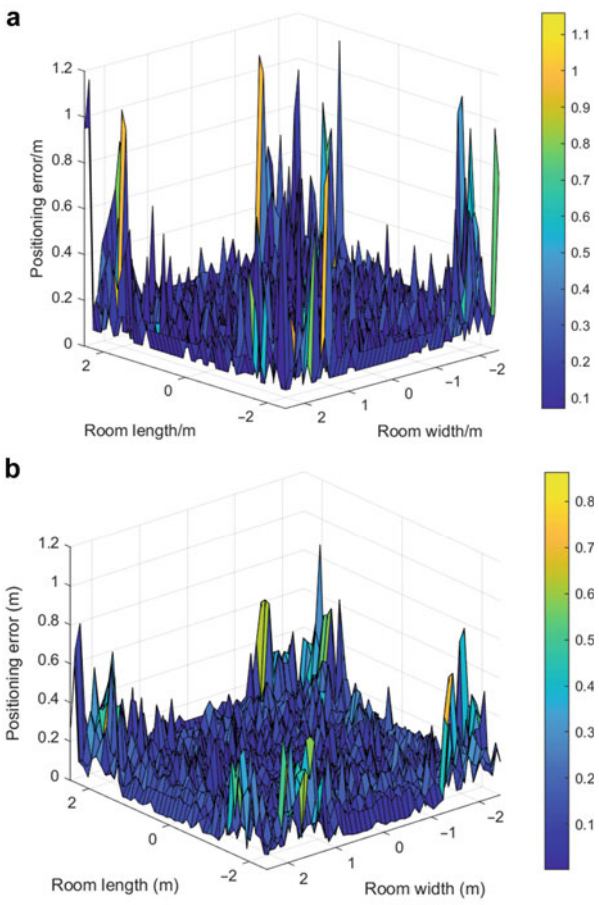


Fig. 41.12 Location error distribution diagram. (**a**) Traditional fingerprint method and (**b**) proposed algorithm

Conclusions

In this study, a novel fusion clustering positioning model using visible light was proposed. By combining k-medoids with the DBSCAN clustering algorithm, the high localization performance of the k-medoids algorithm was used to solve noise problems. This method of clustering followed by positioning is suitable for complex indoor environments.

Coarse positioning improved the positioning efficiency and fine positioning ensured high accuracy, which significantly improved the overall performance. The positioning performance of the whole room was improved by 31% compared with traditional algorithms. The accuracy of this method was verified through simulations and algorithm comparisons.

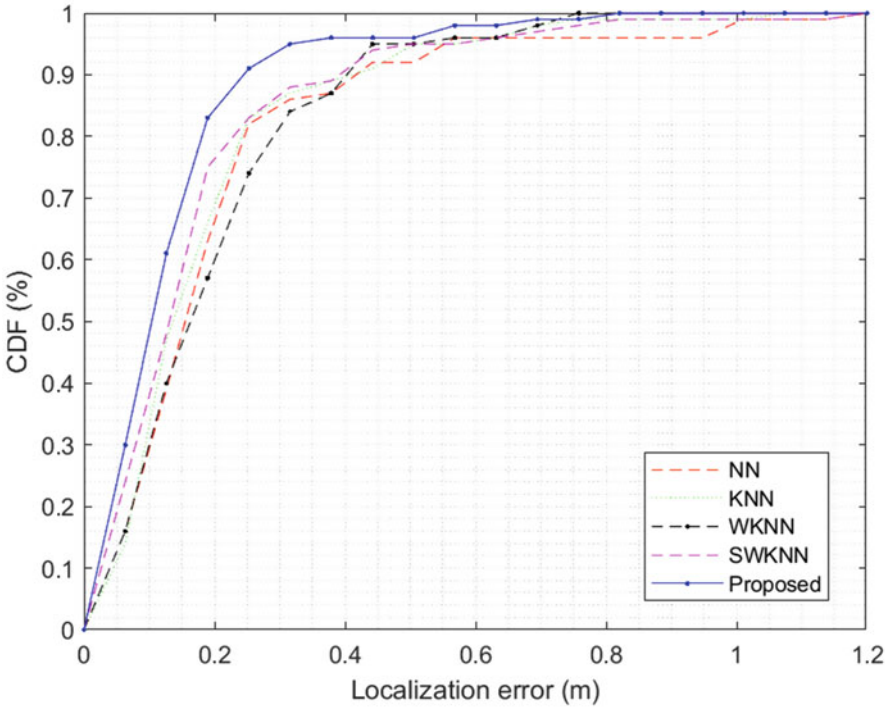


Fig. 41.13 Cumulative probability of errors in different algorithms

Table 41.3 Comparison of positioning accuracy of different positioning methods

Algorithm	Average positioning error (cm)
NN	20.26
KNN	18.51
WKNN	19.87
SWKNN	17.37
Proposed	13.4154

In future work, we should consider an adaptive calibration method to improve the positioning effect of corners and better integrate different positioning technologies to overcome the limitations of a single positioning technology.

References

1. Filippoupolitis, A.; Oliff, W.; Loukas, G. Bluetooth Low Energy Based Occupancy Detection for Emergency Management. In Proceedings of the 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), Granada, Spain, 14–16 December 2016; pp. 31–38.

2. Tekler, Z.D.; Low, R.; Yuen, C.; Blessing, L. Plug-Mate: An Iot-Based Occupancy-Driven Plug Load Management System in Smart Buildings. Build. Environ. 2022, 223, 109472.
3. Balaji, B.; Xu, J.; Nwokafor, A.; Gupta, R.; Agarwal, Y. Sentinel: Occupancy Based Hvac Actuation Using Existing Wifi Infra-structure within Commercial Buildings. In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, Roma, Italy, 11–15 November 2013; p. 17.
4. Tekler, Z.D.; Chong, A. Occupancy Prediction Using Deep Learning Approaches across Multiple Space Types: A Minimum Sensing Strategy. Build. Environ. 2022, 226, 109689.
5. Tekler, Z.D.; Low, R.; Gunay, B.; Andersen, R.K.; Blessing, L. A Scalable Bluetooth Low Energy Approach to Identify Occupancy Patterns and Profiles in Office Spaces. Build. Environ. 2020, 171, 106681.
6. Yousif, B.B.; Elsayed, E.E.; Alzalabani, M.M. Atmospheric Turbulence Mitigation Using Spatial Mode Multiplexing and Modified Pulse Position Modulation in Hybrid RF/FSO Orbital-Angular-Momentum Multiplexed Based on Mimo Wireless Communications System. Opt. Commun. 2019, 436, 197–208.
7. Sejan, M.A.S.; Rahman, M.H.; Aziz, M.A.; Kim, D.-S.; You, Y.-H.; Song, H.-K. A Comprehensive Survey on Mimo Visible Light Communication: Current Research, Machine Learning and Future Trends. Sensors 2023, 23, 739.
8. Ke, X.; Ding, D. Wireless Optical Communication, 2nd ed.; Science Press: Beijing, China, 2022.
9. Gu, W.; Aminikashani, M.; Deng, P.; Kavehrad, M. Impact of multipath reflections on the performance of indoor visible light positioning systems. J. Light. Technol. 2016, 34, 2578–2587.
10. Saadi, M.; Zhao, Y.; Wuttisttikulkij, L.; Khan, M.T.A. A heuristic approach to indoor localization using light emitting diodes. J. Theor. Appl. Inf. Technol. 2016, 84, 332–338.
11. Saadi, M.; Ahmad, T.; Zhao, Y.; Wuttisttikulkij, L. An LED Based Indoor Localization System Using k-Means Clustering. In Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016; pp. 246–252.
12. Wang, H. Research on Positioning Method of Indoor Visible Light Based on Position Fingerprint. Master's Thesis, Xi'an University of Electronic Science and Technology, Xi'an, China, 2018.
13. Liu, R.; Liang, Z.; Yang, K.; Li, W. Machine learning based visible light indoor positioning with single-LED and single rotatable photo detector. IEEE Photonics J. 2022, 14, 1–11.
14. Martínez-Ciro, R.A.; López-Giraldo, F.E.; Luna-Rivera, J.M.; Ramírez-Aguilera, A.M. An Indoor Visible Light Positioning System for Multi-Cell Networks. Photonics 2022, 9, 146.
15. Long, Q.; Zhang, J.; Cao, L.; Wang, W. Indoor Visible Light Positioning System Based on Point Classification Using Artificial Intelligence Algorithms. Sensors 2023, 23, 5224.
16. Komine, T.; Nakagawa, M. Fundamental analysis for visible-light communication system using LED lights. IEEE Trans. Consum. Electron. 2004, 50, 100–107.
17. Ding, D.Q.; Ke, X.Z.; Li, J.X. Design and simulation on the layout of lighting for VLC system. Opto-Electr. Eng. 2007, 34, 131–134.
18. Van, M.T.; Van Tuan, N.; Son, T.T.; Le-Minh, H.; Burton, A. Weighted k-nearest neighbour model for indoor VLC positioning. IET Commun. 2017, 11, 864–871.
19. Maheepala, M.; Kouzani, A.Z.; Joordens, M.A. Light-based indoor positioning systems: A review. IEEE Sensors J. 2020, 20, 3971–3995.
20. Xu, Y.; Wang, X. Indoor positioning algorithm of subregional visible light based on multilayer ELM. J. Hunan Univ. Nat. Sci. 2019, 46, 125–132.
21. Zheng, J.; Li, K.; Zhang, X. Wi-Fi Fingerprint-Based Indoor Localization Method via Standard Particle Swarm Optimization. Sensors 2022, 22, 5051.
22. Wang, K.; Yu, X.; Xiong, Q.; Zhu, Q.; Lu, W.; Huang, Y.; Zhao, L. Learning to improve WLAN indoor positioning accuracy based on DBSCAN-KRF algorithm from RSS fingerprint data. IEEE Access 2019, 7, 72308–72315.

23. Tran, H.Q.; Ha, C. Improved visible light-based indoor positioning system using machine learning classification and regression. Appl. Sci. 2019, 9, 1048.
24. Saadi, M.; Saeed, Z.; Ahmad, T.; Saleem, M.K.; Wuttisittikulkij, L. Visible light-based indoor localization using k-means clustering and linear regression. Trans. Emerg. Telecommun. Technol. 2019, 30, e3480.
25. Koçoğlu, F.Ö. Research on the success of unsupervised learning algorithms in indoor location prediction. Int. Adv. Res. Eng. J. 2022, 6, 148–153.
26. Tao, Z.; Song, Q.; Jin, X. WLAN indoor localization algorithm based on fast K-medoids clustering. Electr. Des. Eng. 2017, 25, 109–113.
27. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD-96 Proc. 1996, 96, 226–231.
28. Wang, Y.; Gao, X.; Dai, X.; Xia, Y.; Hou, B. WiFi Indoor Location Based on Area Segmentation. Sensors 2022, 22, 7920.
29. Liu, Y.; Yu, X.; Xie, S.; Liu, S.; Zhu, P. Channel state information localization based on improved DBSCAN clustering algorithm. Electr. Meas. Technol. 2022, 45, 169–173. 22.
30. Gradim, A.; Fonseca, P.; Alves, L.N.; Mohamed, R.E. On the Usage of Machine Learning Techniques to Improve Position Accuracy in Visible Light Positioning Systems. In Proceedings of the 2018 11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), Budapest, Hungary, 18–20 July 2018; pp. 1–6.
31. Tran, H.Q.; Ha, C. Machine learning in indoor visible light positioning systems: A review. Neurocomputing 2022, 491, 117–131.
32. Tran, H.Q.; Ha, C. High precision weighted optimum K-nearest neighbors algorithm for indoor visible light positioning ap-plications. IEEE Access 2020, 8, 114597–114607.

Experimental Study on Chromaticity Control in Visible-Light Communication Systems

42

Contents

Introduction 1550
Multicolor VLC-System Model and Foreign and Domestic Research Progress 1551
System Model 1551
Domestic Research Progress 1552
Foreign Research Progress 1555
Theoretical Basis 1557
Research on RGB LED Dimming 1558
Dimming and Calculation Methods 1558
Simulation Analysis of Dimming 1560
Dimming Experiment 1564
Design and Experiment of an RGB-LED Visible-Light Communication System 1571
System Working Principle and Experimental Device 1571
Analysis of Experimental Results 1575
Conclusion 1577
References 1579

Abstract

To apply visible-light communication (VLC) systems in different scenarios, this study utilizes a temperature-adjustable light source equipped with RGB light-emitting diodes (LEDs) and applies it to a visible-light communication system. It utilizes color-division multiplexing technology to achieve three-channel communication, thereby improving the communication bandwidth of the system. The communication system adopts three constant-current driving circuits to control the duty cycle of the pulse-width modulation (PWM) of each channel, thereby changing the proportion of RGB LEDs and obtaining different color temperatures to control the color for mixed-color LEDs. The experimental results show that when adjusting the color temperature, the change in luminous flux is very small, with fluctuations of less than 2.24%. When adjusting the brightness, the color-temperature fluctuation is within 40 K, which is less than the 50-K color-temperature limit that the human eye can distinguish, and the average color-temperature

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_42

error is 0.609%. A color tolerance less than 5.5×10^{-3} indicates a good dimming effect. The communication performance of the system is better in the high color-temperature range, which is significantly superior to the low color-temperature range. When the error rate is below 3.8×10^{-3}, the total modulation bandwidth of the three channels reaches 11.7 MHz.

Keywords

Adjustable color temperature · Pulse width dimming · Color division multiplexing · Visible light communication

Introduction

When a white light-emitting diode (LED) is used for the dual functions of lighting and communication, it is important to consider the impact of the color of the light source on the human eye, in addition to considering its communication performance. The lighting environment can have positive or negative psychological, physiological, and visual impacts [1, 2]. After a long period of natural selection, the continuous dynamic change in the color temperature of a natural light source feels the most comfortable to human beings. Color temperature is a physical quantity used in lighting optics to define the color of a light source. When a blackbody is heated to a temperature where the color of the emitted light matches that of a light source, the temperature at which the blackbody is heated is called the color temperature of the light source, also known as the color temperature.

The color temperature of sunlight is approximately 2500 K at sunrise. The color temperature rises to approximately 3500 K 1 h after sunrise and 5300–6500 K at midday. The color temperature then changes from high to low, reaching approximately 2800 K at sunset. The color temperature of the light source obviously affects people's cognition of the visual conditions around them. It plays a role in treating the symptoms of insomnia and alleviating the effect of jet lag, and has a significant impact on people's work efficiency [3].

The color temperature of the light source has a significant impact on the human nerve center [4], and cold light sources are prone to stimulate the nervous system. In a high color-temperature environment, the human heart rate increases, pulse rate accelerates, blood pressure rises, and the mind feels exuberant. In contrast, warm-colored light sources soothe emotions, reduce pressure, and lower the heart rhythm and pulse rate [5]. In terms of psychological perception, low-color-temperature light sources (warm light) make people feel warm, whereas high-color-temperature white light (cool light) makes people feel cold [6].

The light-source chromaticity must be quantitatively controlled to dynamically change the LED color temperature. It is important to study LED light sources with adjustable color temperatures in visible-light communication systems to adjust the color temperature of the LEDs according to the environment to create a healthy light environment in which people feel comfortable. In general, the color-temperature

range of indoor lighting is 3000–6500 K [7]. Light sources with different color temperatures are used on different occasions to meet people's psychological and physiological needs. Therefore, adjusting the color temperature of the light source in a visible-light communication system and controlling its chromaticity is a hot topic in current research.

Visible-light communication (VLC) in an indoor environment must not only provide communication functions, but also meet the lighting needs. Some areas, such as bars and photography studios, not only require the light source to be white, but may also require the color and chromaticity to be adjustable. Therefore, controlling the mixing and chromaticity of a multicolor transmission system should also be considered for VLC light sources.

In this study, controlling the chromaticity of a light source in a visible-light communication system of red-green-blue (RGB) LEDs was investigated and analyzed. By controlling the duty cycle of each channel's pulse-width modulation (PWM), and thereby changing the ratio of red, green, and blue LEDs, different color temperatures can be obtained to control the chromaticity of mixed-color LEDs. They can then be combined with existing communication technology to for communication.

Multicolor VLC-System Model and Foreign and Domestic Research Progress

System Model

A traditional visible-light communication system model mainly includes a sender, a free-space transmission channel, and a receiver. A diagram of a common visible-light communication system is shown in Fig. 42.1. The visible-light communication process can be described as follows: First, the signal generator generates the signal to be sent, passes it through the driving circuit to cause the LED light source to emit light, and modulates the generated signal onto the LED light source. After passing through the indoor free-space channel, it reaches the receiving end. The photodetector at the receiving end detects the received optical signal and converts it into an electrical signal, which is then processed by the receiving circuit and displayed on an oscilloscope.

The principles of baseband modulation and photoelectric signal transformation in visible-light communication systems involve loading the original signal onto the

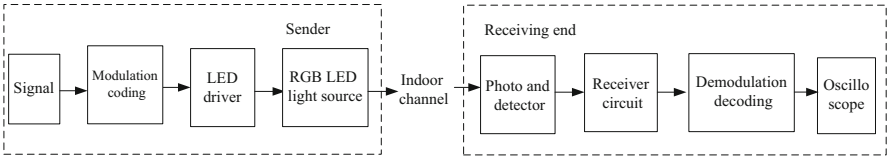


Fig. 42.1 Block diagram of a visible-light communication system

optical signal through an LED driver module after baseband modulation and digital-to-analog transformation. Then, the signal is transmitted from the sending end using the high-speed flashing changes of the LED light and transmitted through indoor channels to reach the receiving end of the system.

The driving circuit of an LED generally includes signal-amplification and DC-bias modules. The signal-amplification module appropriately amplifies the baseband signal, based on the dynamic range and communication requirements of the LED. The DC bias adds a DC component to the AC signal, allowing the LED to operate in a linear region. At the sending end, the electrical signals are converted into optical signals, whereas at the receiving end, photoelectric detectors are used to convert the optical signals into electrical signals. The receiving circuit filters, amplifies, reshapes, and demodulates the converted electrical signal before outputting it.

When a visible-light communication system has specific requirements for the transmission distance of a free-space channel, optical lenses may need to be added to focus the optical path to improve the signal-to-noise ratio (SNR) of the system.

Domestic Research Progress

In research on visible-light communication systems, fluorescent-powder LEDs that met both daily lighting needs and communication-system functions were first used. They had good performance and a simple structure; therefore, they were favored in the early research stages. However, phosphor-based LEDs have a slow response time, which limits the modulation bandwidth of the devices. Therefore, in recent years, research on high-speed visible-light communication has tended to use RGB LEDs as emission light sources. RGB LEDs have a faster response speed and a larger modulation bandwidth, and provide multiple optical-communication channels with different wavelengths. Color-division multiplexing can be used to superpose the modulation bandwidth.

In the visible-light communication field, an increasing number of scholars are studying multicolor visible-light communication systems to improve their performance through various hardware-improvement design schemes and modulation-coding techniques. In 2006, Ke et al. [8] investigated visible-light communication and its key technologies. The following year, the team [9] designed and simulated the light-source layout of a visible-light communication system and designed and verified the optimal layout of the four light sources.

To improve the communication speed, visible-light communication systems have developed from single-channel communication to multi-channel communication. In 2013, Wang et al. [10] used RGB LEDs and phosphor-based LEDs as emitting light sources, and conducted multiplexing experiments using wavelength-division multiplexing (WDM). The experimental results showed that when the communication distance was 66 cm and the BER was lower than 3.8×10^{-3}, a visible-light communication transmission with a 225-Mbps uplink communication rate and 575-Mbps downlink communication rate was achieved.

In the same year, Wu et al. [11] established a visible-light communication system using commercial RGB LEDs mixed with white light as the emission light source, using carrier-free amplitude-phase modulation and wavelength-division multiplexing technology. They successfully conducted visible-light communication experiments with a 25-cm communication distance and 3.22-Gbps data-transmission rate for the three channels, and an error rate of less than 10^{-3}.

In 2015, Wang et al. [12] established a high-speed visible-light communication system using red-green-blue-yellow (RGBY) LEDs as light sources, high-order carrier-free amplitude phase modulation, and mixed equalizers. The experimental results showed that the total communication rate of the three channels was 8 Gbps, the error rate was less than 3.8×10^{-3}, the communication distance was 1 m, and the illumination value of the receiving surface was 450 lx.

In the same year, Huang et al. [13] successfully demonstrated a gigabit-per-second visible-light communication transmission experiment based on RGB LEDs in free space at an 80-cm communication distance. Single-carrier 64-way quadrature-amplitude modulation (64QAM), 64QAM-orthogonal frequency-division multiplexing (OFDM), and bit-power loaded OFDM (BPL-OFDM) signals were demonstrated on an 80-cm free-space transmission at communication rates of 1.05 Gbps, 1.05 Gbps, and 1.42 Gbps, respectively, with error rates less than 3.8×10^{-3}. Compared with systems without equalizers, the bit-error rate performance improved by at least one order of magnitude. For BPL-OFDM modulation signals, the overall data rate increased from 1.13 Gbps to 1.42 Gbps by deploying equalization circuits.

In 2015, Ke et al. [14] constructed an RGB-LED visible-light communication video-transmission system and conducted experiments on video data-stream transmission with a 2-m communication distance and 1.5-Mbps data-transmission rate.

In 2016, Zhang et al. [15] considered the impact of color combinations on the communication rate based on RGB LEDs and increased the transmission rate to 10.4 Gbps using a rate-adaptive OFDM modulation method. In the same year, Chi et al. [16] proposed a new phase-shifted (PS) Manchester code based on eight-way pulse-amplitude modulation (PAM-8), and their experiments proved that this code could reduce noise and distortion. In a visible-light communication system based on RGB LEDs, with the help of PS Manchester encoding and wavelength-division multiplexing, information was transmitted over a 1-m communication distance, with a 3.375-Gbps total data rate and an error rate of less than 3.8×10^{-3}.

In 2017, Chen et al. [17] proposed a new Manchester-coded PAM-8 modulation scheme based on a hybrid time–frequency domain equalization scheme and applied it to a visible-light communication system based on RGB LEDs. The experimental results indicated that the overall data rate reached 4.05 Gbps. In the same year, Lu et al. [18] established a visible-light communication system using RGB LEDs combined with OFDM modulation. The experimental results showed that when the link was in single-input single-output mode, the system achieved a total communication rate of 3.48 Gbps. When the link was in multiple-input multiple-output mode, a 6.36-Gbps communication rate, 3.8×10^{-3} bit-error rate, and 1-m communication distance were successfully achieved.

In 2017, Kong et al. [19] implemented a wavelength-division multiplexing visible-light communication system based on RGB LEDs using 32 orthogonal-amplitude mapping (OAM) OFDM modulation in visible-light communication experiments. The experimental results showed that the data-transmission rates achieved by the red, green, and blue channels in a 10-m underwater channel were 4.17 Gbps, 4.17 Gbps, and 1.17 Gbps, respectively.

In 2018, Zhu et al. [20] used red-green-blue-cyan-yellow (RGBCY) LEDs as emission light sources and conducted experiments on indoor visible-light communication systems using wavelength-division multiplexing. The experimental results showed that at a 1-m communication distance, the total system communication rate was 10.72 Gbps, and the error rate was less than 3.8×10^{-3}.

In 2019, Deng et al. [21] proposed a multipulse position modulation (MPPM) dimming control scheme that supported different dimming target values. The proposed scheme improved the throughput and transmission efficiency of visible-light communication systems. In the same year, Zuo et al. [22] considered reliable data transmission and lighting quality when transmitting signals in visible-light communication systems and proposed a block-coding-based dimming scheme to construct a codeword set. The dimming was controlled by changing the ratio of the two levels through modulation using a key control. The experimental results showed that under different dimming levels, the scheme maintained good error performance and a constant transmission efficiency. They conducted communication experiments with a 2.5-m communication distance, 2.3-Mbps transmission rate, and a bit-error rate of less than 10^{-4}.

In 2020, Hu et al. [23] proposed a WDM visible-light communication transmission system based on 4×4 LED arrays using a discrete multi-tone (DMT) modulation technique and achieved a total communication rate of 20.09 Gbps with almost error-free communication experiments over a 1.2-m underwater link. In addition, for future LED-based WDM visible-light communication systems, the authors proposed a silicon-based multi-color LED array chip and a feasible filter scheme.

In 2022, Zhang [24] used white LEDs as the emission light source in a visible-light communication system using on–off keying (OOK) modulation. At a 3-m communication distance and 56-kbps data-transmission rate, he completed communication experiments that achieved both lighting effects and low-bit-error-rate signal transmissions. In the same year, Qiu et al. [25] successfully achieved a 25.2-Gbps communication rate and an almost error-free visible-light communication experiment with three LEDs and five micro-LEDs, using WDM technology at a 25-cm communication distance.

In 2023, Min [26] used color shift-keying modulation technology in a multi-color LED communication system. By testing the frequency response and high-speed communication performance of a four-color visible-light communication system, a single-color analog bandwidth on a 120-MHz communication system was achieved. The communication performance under various lighting indicators was analyzed, and a communication experiment with a communication rate of 50 Mbps and good lighting performance was conducted.

Foreign Research Progress

To prove that LEDs could simultaneously serve both lighting and communication functions, scholars from the Nakagawa Laboratory in Japan conducted a theoretical analysis and system simulation of visible-light communication system channels in 2000. Two years later, researchers conducted further studies on the characteristics, channel model, noise model, signal-to-noise ratio (SNR), and other performance indicators of LED light sources [27]. Some unknown interference may exist, owing to the free space of the visible-light communication system channel. Modulation technology can be used to improve the system's interference resistance.

In 2010, Vucic et al. [28] implemented a visible-light wireless point-to-point communication link with a total transmission rate of 513 Mbps, based on commercial thin-film high-power phosphorescent white LEDs and orthogonal amplitude modulation. Additionally, obstacles to improving data rates and potential remedial measures were discussed. If these obstacles were overcome, data rates exceeding 1 Gbps could be achieved under the same settings and experimental conditions.

In 2011, Kottke et al. [29] utilized RGB LEDs as emission light sources in visible-light communication systems using WDM and DMT modulation. At a 12-cm communication distance and a 1000-lx light intensity, they achieved communication rates of 293.7 Mbps, 232.4 Mbps, and 286 Mbps for three channels, with a total communication rate of 803 Mbps and an error rate of less than 2×10^{-3}.

The following year, the team [30] utilized RGB LEDs to mix white light and achieved communication rates of 376 Mbps, 439 Mbps, and 430 Mbps for the red, green, and blue channels, respectively, based on WDM and DMT modulation. The total communication rate was 1.25 Gbps, and the communication distance was 10 cm. Moreover, the illumination value at the receiving end met the working-environment indicator of 1000 lx, and the error rates of all three channels were below 10^{-3}.

In 2012, Cossu et al. [31] achieved a 1.5-Gbps single-channel communication rate on commercial RGB white LEDs using multicarrier modulation. Combined with WDM, the total system communication rate under illumination reached 3.4 Gbps at a 30-cm communication distance, and the system's BER was less than 2×10^{-3}.

In 2014, the team [32] again implemented a bidirectional visible-light communication system using WDM. The experimental results showed that when the communication distance varied from 1.5 m to 4 m, the uplink communication rate of the system varied in the range of 1.1 Gbps to 1.5 Gbps, and the range of the downlink communication rate varied in the range of 5 Gbps to 5.6 Gbps. The range of the light intensity ranged from 120 lx to 720 lx, and the BERs were all less than 3.8×10^{-3}. In the next year, the team [33] achieved four-channel information transmission in a visible-light communication system, when the BER of each channel was less than 3.8×10^{-3}. They achieved a data rate greater than 5 Gbps for the downlink communication and greater than 1.5 Gbps for the uplink communication over a transmission distance of 1.5–4 m indoors.

In 2014, Sung et al. [34] proposed an RGB-LED dimming scheme based on DMT modulation that could control the color in a high-speed visible-light communication

system. The authors conducted experiments and analyses under various modulation conditions and transmission distances. The experimental results showed that at a 30-cm transmission distance and a 512-point fast Fourier transform, a total transmission rate of 103.5 Mbps for three channels was achieved, with an error rate of less than 3.8×10^{-3}.

In 2016, Lee et al. [35] optimized and improved the traditional variable pulse-position modulation (PPM), ensuring that LEDs could stably transmit data, even in environments with rapid changes in light intensity. A visible-light communication system test bench with rapidly changing brightness and variable PPM was designed. The system light source was composed of 30 white LED chips, achieving a communication rate of 200 kbps at a 6-m communication distance, and bit-error rate of 6×10^{-3} in communication experiments.

In 2017, Sewaiwar et al. [36] proposed a high-performance time-division multiplexing-based multiuser multiple-input multiple-output (MIMO) system for efficient indoor visible-light communication. The transmitting light source in this system used RGB LEDs for parallel data transmissions. Simulation results showed that the proposed scheme provided an approximately 6-dB gain in the signal-to-noise ratio at a BER of 3×10^{-5} compared to a conventional single-input single-output system. In addition, the data rate reached 66.7 Mbps at a BER of 10^{-3}.

In 2018, Bian et al. [37] used RGB LEDs mixed with white light as the light source and used OFDM and WDM techniques in visible-light communication systems. The communication distance was 50 cm, the data-transmission rate was 10.2 Gbps, and the error rate was less than 3.8×10^{-3} in high-speed visible-light communication experiments. The following year, the team [38] used four different colored LEDs as emission light sources in a visible-light communication system and used OFDM modulation to complete a VLC transmission experiment with a communication rate of 15.73 Gbps and almost no error codes on a 1.6-m communication link.

In 2019, Jain et al. [39] used RGB LEDs as the emission light source for visible-light communication systems using a color shift-keying modulation method. They proposed a new adaptive nonlinear equalizer based on multivariate-kernel least squares to alleviate the intersymbol interference and nonlinearity of RGB LEDs. The simulation results showed that the system had a lower symbol error rate than existing technologies.

In 2020, Viñals et al. [40] designed a linear precoder for the downlink of visible-light communication systems that simultaneously served multiple users. The emission light source of this system used RGB LEDs to modulate three independent data streams on the three RGB channels and used multilevel PAM to conduct communication experiments with a modulation bandwidth of 15 MHz and a bit-error rate below 10^{-6}.

Multichannel visible-light communication systems using multiple monochromatic LEDs have achieved high communication rates; however, no color-control mechanism exists. The color of the light emitted by multiple LEDs poses a threat to the human eye, and many scholars have investigated this [41, 42].

Theoretical Basis

To analyze the LED color rendering, the Gaussian mathematical model (Y-model) proposed by Ohno et al. is in good agreement with the actual LED spectral distribution and has been widely recognized. This model can be represented as [43]

$$S_{\mathrm{LED}}(\lambda, \lambda_0, \Delta\lambda_{0.5}) = Y_c \frac{g(\lambda, \lambda_0, \Delta\lambda_{0.5}) + 2g^5(\lambda, \lambda_0, \Delta\lambda_{0.5})}{3} \tag{42.1}$$

where $g(\lambda, \lambda_0, \Delta\lambda_{0.5}) = \exp\left[-((\lambda - \lambda_0)/\Delta\lambda_{0.5})^2\right]$; λ is the wavelength; Y_c is the spectral power at the peak wavelength λ_0; and $\Delta\lambda_{0.5}$ is the half peak width.

The tristimulus value of a light source refers to its basic optical characteristics, which are typically used to describe and measure its color. According to the spectral mathematical model of the LED, the tristimulus values X, Y, and Z of the light source to be tested can be calculated using the following formula [44]:

$$\begin{cases} X = K_m \int_{380}^{780} S(\lambda)\bar{x}(\lambda)d\lambda \\ Y = K_m \int_{380}^{780} S(\lambda)\bar{y}(\lambda)d\lambda \\ Z = K_m \int_{380}^{780} S(\lambda)\bar{z}(\lambda)d\lambda \end{cases} \tag{42.2}$$

where $K_m = 683lm/W$ is a constant; $S(\lambda)$ is the relative spectral power distribution; $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are the color-matching function, which is the average data of the CIE 1931 Standard Observer, also known as the spectrum triple-stimulation function [44].

In the actual calculation, the CIE 1931 standard-chromaticity tristimulus value of the light source was obtained using discrete summation and approximate integration calculations. The expression can be expressed as

$$\begin{cases} X = K_m \sum_{380}^{780} S(\lambda)\bar{x}(\lambda)\Delta\lambda \\ Y = K_m \sum_{380}^{780} S(\lambda)\bar{y}(\lambda)\Delta\lambda \\ Z = K_m \sum_{380}^{780} S(\lambda)\bar{z}(\lambda)\Delta\lambda \end{cases} \tag{42.3}$$

where $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ can be determined using lookup tables, and a spectral sampling interval of $\Delta\lambda = 1$ is often used.

The corresponding color coordinates can be determined according to the three obtained light-source stimulus values. After normalizing X, Y, and Z, the color coordinates x, y, and z of the light source can be obtained as follows:

$$\begin{cases} x = \dfrac{X}{X+Y+Z} \\ y = \dfrac{Y}{X+Y+Z} \\ z = \dfrac{Z}{X+Y+Z} \end{cases} \tag{42.4}$$

In lighting systems, the color temperature is a physical quantity used to define the chromaticity of a light source, and can indicate its lighting quality. The relevant color temperature of the light source can be calculated from the chromaticity coordinates; the formula for calculating the color temperature is as follows [45]:

$$T = -437n^3 + 3601n^2 - 6861n + 5514 \tag{42.5}$$

where $n = \frac{x-0.3320}{y-0.1858}$ and x, y is the color coordinate.

The color index is a parameter used to evaluate the degree to which a light source restores the true color of an object. It is generally believed that, under natural light, the color presented by objects is the most authentic, with a color index of 100. The color index of the light source for a standard sample can be expressed as [46]

$$R_i = 100 - 4.6\Delta E_i \, (i = 1, 2, 3, \cdots, 14) \tag{42.6}$$

where ΔE_i is the color difference between the sample under the reference light source and the light source to be tested, and can be obtained from the chromaticity coordinates.

The average color index of the light source for color samples 1–8 is called the general color index R_a:

$$R_a = \frac{1}{8}\sum_{i=1}^{8} R_i \tag{42.7}$$

Research on RGB LED Dimming

Dimming and Calculation Methods

Dimming technology has become an important aspect of LED driving, and three LED dimming methods are commonly used: DC, PWM, and thyristor dimming [47]. When the current is too small, the thyristor will disconnect, making the dimming effect unstable and unsuitable for use in visible-light communication systems [47]. DC dimming is achieved by changing the driving peak current of the LED to alter the luminous flux. This dimming method can affect the chromaticity shift and color-temperature change of the LED [48].

PWM dimming is achieved by changing the duty cycle of the LED light source, thereby changing the average luminous flux. The dimming performance is flexible

and can accurately control the luminous flux, thereby accurately controlling the brightness change of the LED and reducing the effect of current changes during amplitude modulation on the stability of the light source [49]. In engineering practice, color differences caused by PWM dimming are often ignored [50].

In the PWM-dimming method, the peak value of the driving current remains unchanged and the LED is dimmed by adjusting the duty cycle of the PWM. The expression is as follows:

$$I_{\text{avg}} = I_{\text{peak}} \cdot D \tag{42.8}$$

where I_{avg} represents the effective value of the driving current, I_{peak} represents the magnitude of the driving peak current, and D represents the PWM duty cycle.

For the convenience of calculation, it is usually assumed that the PWM duty cycle of the input driver module is directly proportional to the luminous flux output of the light source [51], which satisfies

$$\Phi = D\Phi_{\text{Max}} \tag{42.9}$$

where D is the duty cycle of the input PWM signal, $D < 1$; Φ_{Max} is the maximum output luminous flux; and Φ is the luminous flux of the output signal.

Glassman's law of color mixing provides a three-color mixing-output luminous-flux formula [52]:

$$\Phi = D_r\Phi_r + D_g\Phi_g + D_b\Phi_b \tag{42.10}$$

where D_r, D_g, and D_b are the input duty cycles of the PWM; Φ_r, Φ_g, and Φ_b are the maximum output luminous fluxes of each light source; and Φ is the output luminous flux of the mixed light.

The CIE-1931 standard colorimetric system specifies that the stimulus value Y is equal to the luminous flux; therefore, in the following text, it can be changed to Y. According to the principle of color mixing and the CIE 1931 color-coordinate calculation method, the color coordinates after mixing with the three light sources should satisfy the following [53]:

$$x = \frac{R_r D_r x_r + R_g D_g x_g + R_b D_b x_b}{R_r D_r + R_g D_g + R_b D_b} \tag{42.11}$$

$$y = \frac{R_r D_r y_r + R_g D_g y_g + R_b D_b y_b}{R_r D_r + R_g D_g + R_b D_b} \tag{42.12}$$

where x and y represent the color coordinates of a mixed light source consisting of three types of light sources; (x_r, y_r), (x_g, y_g), and (x_b, y_b) represent the color coordinates of red, green, and blue LEDs, respectively; $R_r = \frac{Y_r}{y_r}$, $R_g = \frac{Y_g}{y_g}$, and $R_b = \frac{Y_b}{y_b}$ represent the sum of the three stimulus values of the light source under full current operation; and D_r, D_g, and D_b represent the PWM input duty cycles of red, green, and blue LEDs, respectively.

From Eqs. (42.10), (42.11), and (42.12), the duty cycles of the three channels are obtained as follows:

$$D_r = \frac{\left(y_g - y_b\right)(x_b - x) + (y - y_b)\left(x_g - x_b\right)}{\left(y_g - y_b\right)(x_b - x_r) + (y_r - y_b)\left(x_g - x_b\right)} \cdot \frac{y_r Y}{y Y_r} \tag{42.13}$$

$$D_g = \frac{(y_b - y_r)(x_r - x) + (y - y_r)(x_b - x_r)}{(y_b - y_r)\left(x_r - x_g\right) + \left(y_g - y_r\right)(x_b - x_r)} \cdot \frac{y_g Y}{y Y_g} \tag{42.14}$$

$$D_b = \frac{\left(y_g - y_r\right)(x_r - x) + (y - y_r)\left(x_g - x_r\right)}{\left(y_g - y_r\right)(x_r - x_b) + (y_b - y_r)\left(x_g - x_r\right)} \cdot \frac{y_b Y}{y Y_b}, \tag{42.15}$$

where Y_r, Y_g, and Y_b are the respective maximum output luminous fluxes when the input duty cycles of the red, green, and blue light sources are 100%; Y is the output luminous flux of a hybrid light source; (x_r, y_r), (x_g, y_g), and (x_b, y_b) are the color coordinates of red, green, and blue, respectively; and (x, y) is the color coordinate of the mixed light source.

Simulation Analysis of Dimming

The duty cycle is proportional to the luminous flux, which determines the LED luminance. The larger the luminous flux, the brighter the LED luminance. Therefore, the duty cycle is directly proportional to the luminous brightness of the LED. The ratio of RGB LEDs with different brightnesses controlled by the circuit were calculated and simulated. The brightness ratio of a single LED was changed, and the changes in color temperature and color coordinates of the RGB-LED hybrid light source were observed through simulation.

The arrangement of the three types of LEDs is shown in Fig. 42.2, with an observation-surface size of 500 mm × 500 mm and a distance of 100 cm from the emitter. In the simulation, the distance remained unchanged; only the brightness ratio of the LED was changed. The total luminous flux of the LED was 100 lm.

Simulating Color-Coordinate Changes

When the control circuit adjusts the luminous-flux ratio of RGB LEDs, different color temperatures can be obtained. The changes in different color temperatures and corresponding color coordinates were simulated, based on the ratio calculated from Eqs. (42.11) and (42.12). As shown in Fig. 42.3, color coordinate x shows a decreasing trend with increasing color temperature, whereas color coordinate y fluctuated with increasing color temperature. The fluctuation is greater during the low-color-temperature stage and smaller during the high-color-temperature stage.

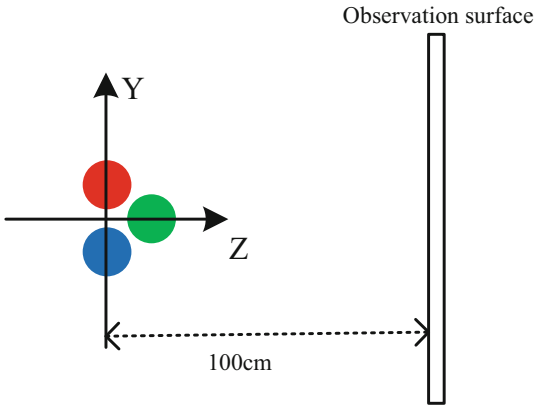


Fig. 42.2 Arrangement of three types of LEDs

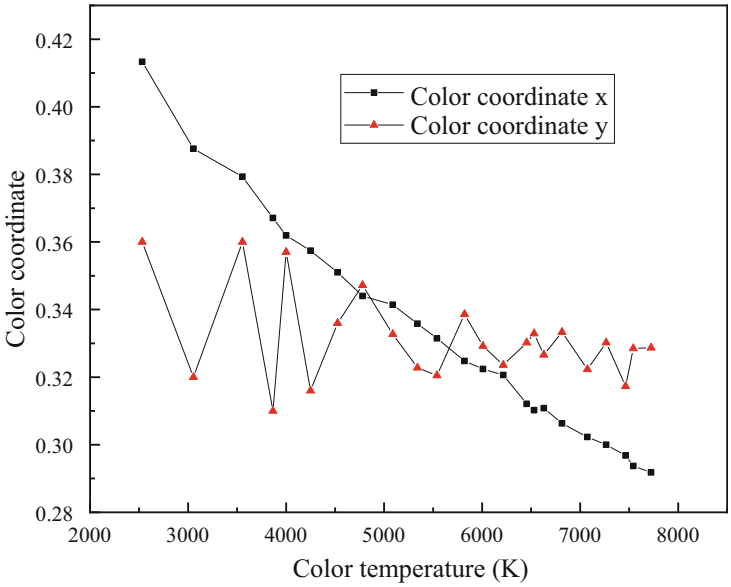


Fig. 42.3 Curves of the color coordinates changing with the color temperature

Figure 42.4 shows the mixed light source on the observation surface when the duty-cycle ratios of the RGB LEDs are set to 0.6: 0.28: 0.12, 0.5: 0.32: 0.19, 0.47: 0.33: 0.2, and 0.45: 0.34: 0.21. At this time, the color temperatures of the mixed light source are 3470 K, 5677 K, 6528 K, and 7265 K, respectively. The color coordinates are (0.4044, 0.3847), (0.3286, 0.3334), (0.3116, 0.3336), and (0.3038, 0.3099), respectively, and the total amount of light is approximately 88 lm.

The simulation showed that the proportions of the RGB LEDs had a significant impact on the color coordinates and temperature of the synthesized white light, with

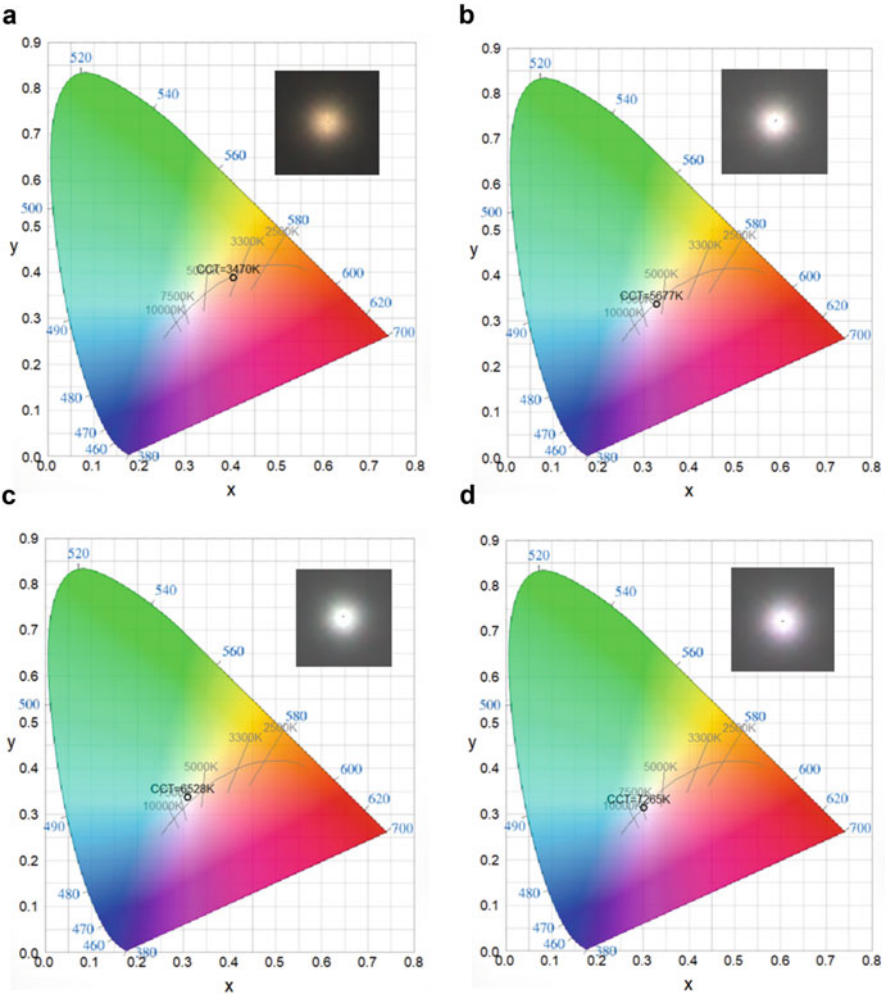


Fig. 42.4 Simulating light sources at different color temperatures. (**a**) 3470 K. (**b**) 5677 K. (**c**) 6528 K. (**d**) 7265 K

a relatively small impact on the luminous flux. By changing the duty cycle of the red, green, and blue LEDs, the hybrid white light source changed from warm white to cold white.

Simulating Color-Temperature Changes

A constant distance of 100 mm was maintained between the receiving surface and light source, and the changes in the average light intensity of the receiving surface under different color temperatures were simulated. As shown in Fig. 42.5, the average light intensity of the receiving surface at different color temperatures was

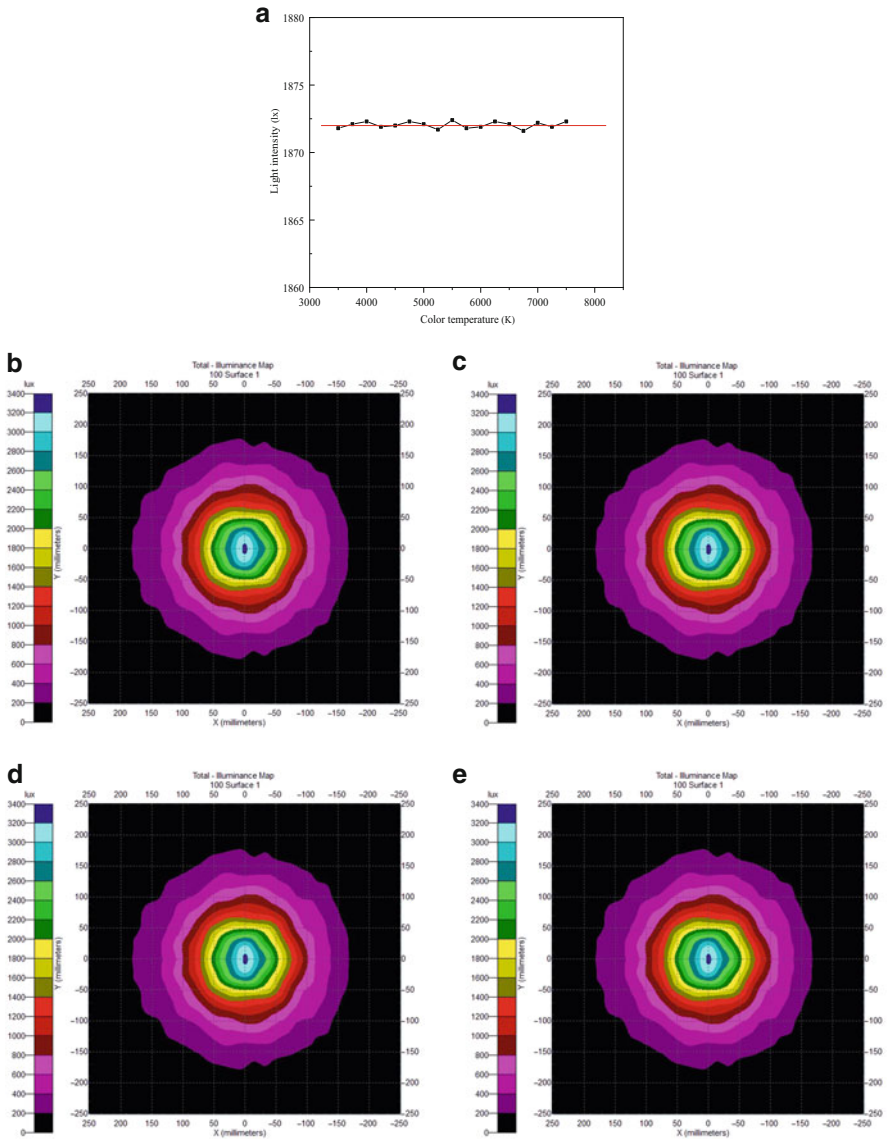


Fig. 42.5 Light intensity distribution under different color temperatures. (**a**) The variation of light intensity with color temperature; (**b**) Color temperature is 3500 K; (**c**) Color temperature is 5500 K; (**d**) Color temperature is 6500 K; and (**e**) Color temperature is 7300 K

approximately 1872 lx, with an error of less than 5 lx between the maximum and minimum values. When the color temperatures were 3500 K, 5500 K, 6500 K, and 7300 K, the simulated average light-intensity distribution of the receiving surface showed almost no difference. In summary, as the color temperature of the light

source changes, the change in the illumination intensity of the receiving surface becomes very small and can be ignored.

Simulating Different Transmission Distances

A color temperature of 6500 K was fixed and the light intensity on the receiving surface at different distances was simulated. As shown in Fig. 42.6, the energy received by the receiving surface was increasingly less dispersed as the distance between the light source and the receiving surface increased. When the distance between the light source and the observation surface was 50, 100, 200, and 300 cm, the simulated light-intensity map of the receiving surface showed that the energy dispersion increased with increasing distance. In summary, as the distance increased, the light intensity on the receiving surface decreased, and the distance had a significant effect on the light intensity on the receiving surface.

Dimming Experiment

This experiment used hexagonal RGB LED beads to form a luminescent light source; an HP350 spectrophotometer produced by Hangzhou Double Color Intelligent Testing Instrument Co., Ltd. was used to measure the light color parameters. Table 42.1 lists the light-color parameters when the three-color LED was operated at full current, where FWM is the full width at half maximum. Figure 42.7 shows the relative spectral power distribution.

The duty cycle of the LED was adjusted and its brightness was measured. Figure 42.8 shows the variation curve of the measured LED brightness with the duty cycle. The brightness of the green light changed significantly with the duty cycle, whereas the changes in the red and blue lights were relatively similar, with relatively small changes compared to the green light.

The dimming device consists of a switching power supply, control system, driving circuit, and red, green, and blue LED light sources. Figure 42.9 shows a block diagram of the dimming device. The control system sends instructions to the driving circuit to drive a three-channel LED, obtaining white light with different color temperatures based on the different duty cycles, thereby adjusting the LED color temperature.

RGB LEDs can be mixed in different proportions to obtain a color-temperature range of 2500–8500 K. To study the optimal color-rendering performance of the light-source module during the dimming process, 12 sets of combinations with good color-rendering performance were selected. The mixed-light parameters measured using an HP350 spectrophotometer are listed in Table 42.2.

From Table 42.2, it can be seen that, compared with the measured color temperature, the color deviation calculated according to Eq. (42.5) is lower than the 50-K color-temperature limit that can be resolved by the human eye. The average color-temperature error was 0.609%, and the difference between the measured and calculated values was small. Therefore, the color-temperature adjustment in the experiment was relatively stable.

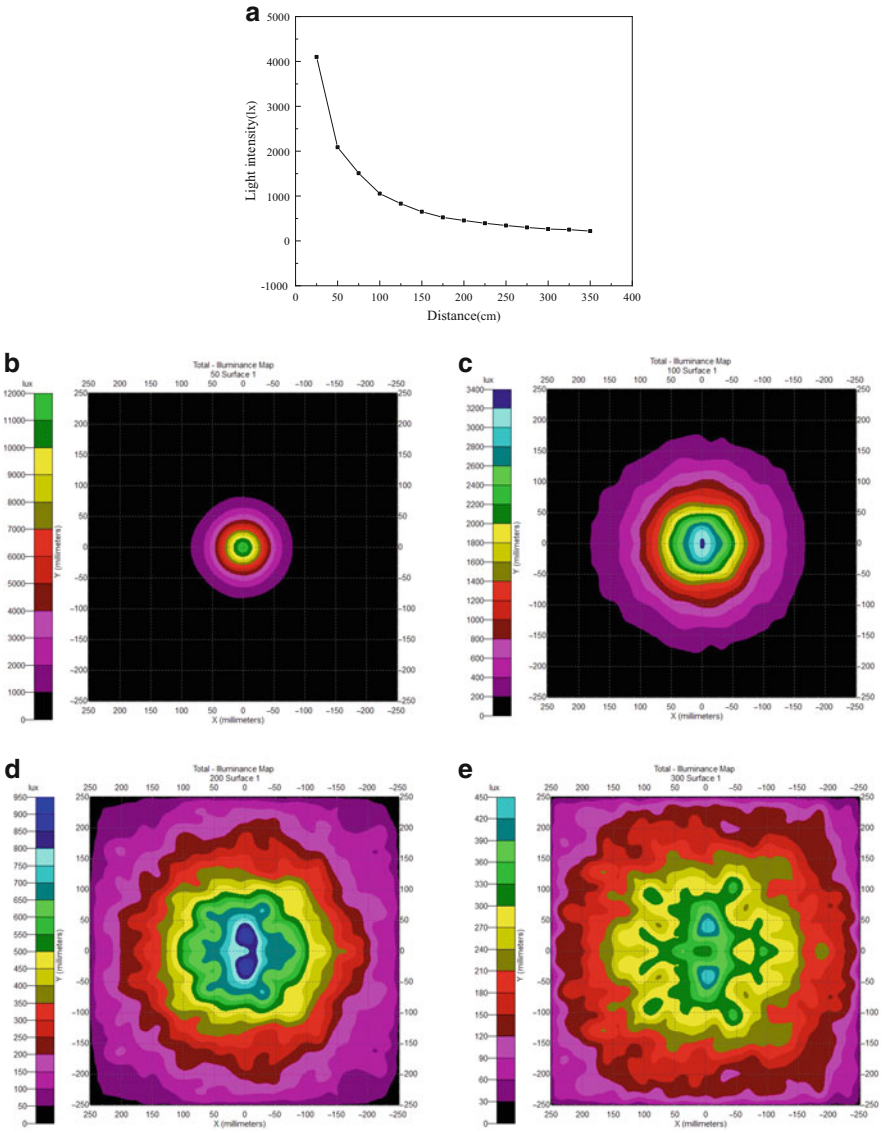


Fig. 42.6 Light intensity distribution at different distances. (**a**) The variation of light intensity with distance; (**b**) Distance is 50 cm; (**c**) Distance is 100 cm; (**d**) Distance is 200 cm; and (**e**) Distance is 300 cm

The theoretical color coordinates of the synthesized light source were obtained from Eqs. (42.3) and (42.4), and the actual color coordinates were measured using a spectrophotometer. Figure 42.10 shows the variation curves of the measured and

Table 42.1 Three LED parameters

LED type	Chromatic coordinates	Peak wavelength (nm)	FWHM (nm)	Luminous intensity (cd)	Luminous flux (lm)
Red	(0.6894, 0.3099)	630	20	1.8–2.0	67.9
Green	(0.1689, 0.7232)	520	20	5.0–6.0	178.1
Blue	(0.1434, 0.0416)	460	20	1.0–1.5	29.0

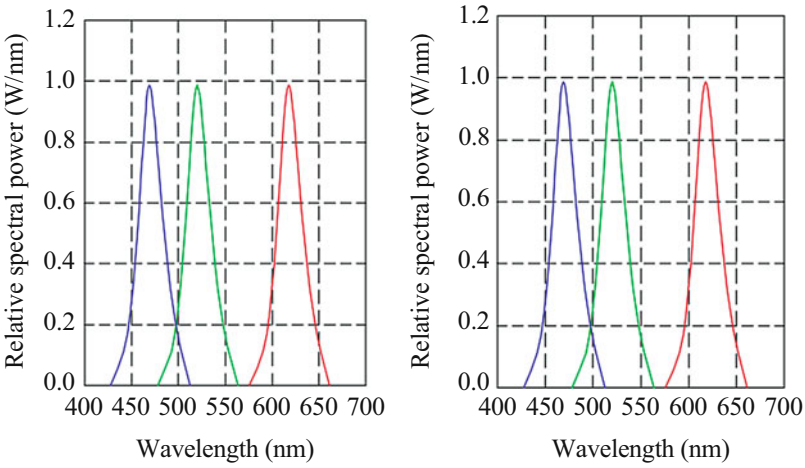


Fig. 42.7 Relative spectral power-distribution curves of three types of LEDs

calculated color coordinates with the color temperature. The error of the calculated color coordinate x was less than 0.0007, and the error of the color coordinate y was less than 0.0014. The measured color coordinates agreed well with the calculated color coordinates. The color coordinates have a significant impact on low-color-temperature light sources, whereas they have a relatively small impact on high-color-temperature light sources, which is consistent with the simulation results mentioned above.

The duty cycles of the RGB LEDs obtained from Eqs. (42.13), (42.14), and (42.15) can produce mixed white light with different color temperatures. As shown in Fig. 42.11, the duty cycles of the red, green, and blue LEDs varied with the color temperature. As shown in Fig. 42.11, as the color temperature increases, the duty cycle of the red LED decreases, whereas those of the green and blue LEDs increase.

An HP350 spectrophotometer was used to measure the luminous flux, irradiance, color tolerance, light intensity, and color index of the mixed light source. Figure 42.12 shows that the maximum luminous flux set in the experiment was 100 lm, and the luminous flux of the mixed light varies between 92 lm and 100 lm. The resulting luminous flux of the mixed white light is relatively high,

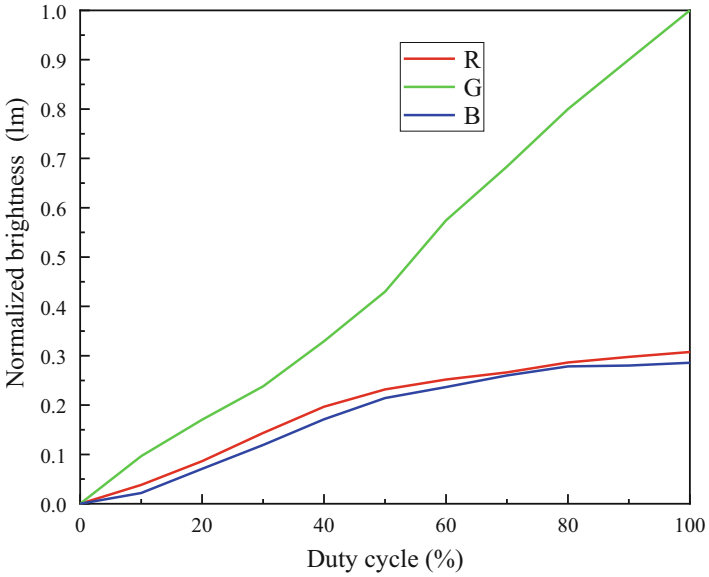


Fig. 42.8 Changes in the LED brightness with the duty cycle

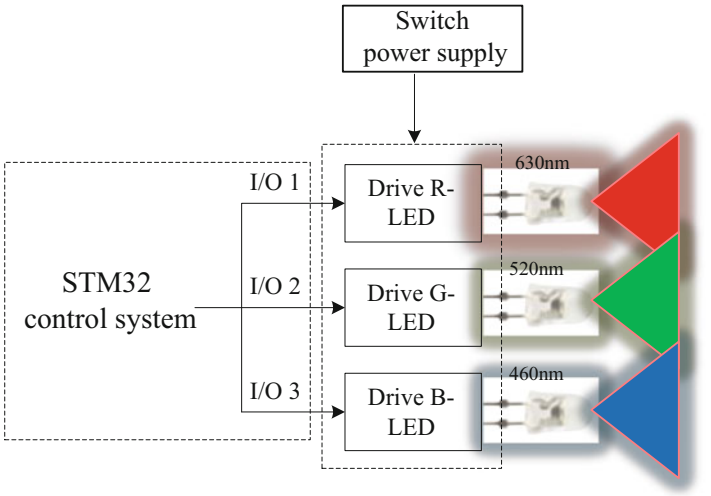


Fig. 42.9 Block diagram of the dimming device

and when the color temperature is adjusted, the fluctuation in the luminous flux is very small, with a fluctuation of 2.24%. Therefore, the influence of the color temperature on the luminous flux was relatively small, which is consistent with the simulation results.

Table 42.2 Comparison between measured and calculated data

RGB proportion	Measured color temperature (K)	Calculated color temperature (K)	Color temperature difference (K)
0.557: 0.301: 0.142	2539	2500	39
0.502: 0.280: 0.218	3030	3000	30
0.476: 0.347: 0.177	3525	3500	25
0.439: 0.304: 0.257	4012	4000	12
0.417: 0.370: 0.213	4532	4500	32
0.405: 0.323: 0.272	5023	5000	23
0.382: 0.361: 0.257	5510	5500	10
0.364: 0.356: 0.281	6032	6000	32
0.347: 0.363: 0.290	6526	6500	26
0.336: 0.355: 0.308	7036	7000	36
0.313: 0.382: 0.305	7540	7500	40
0.302: 0.381: 0.317	8032	8000	32

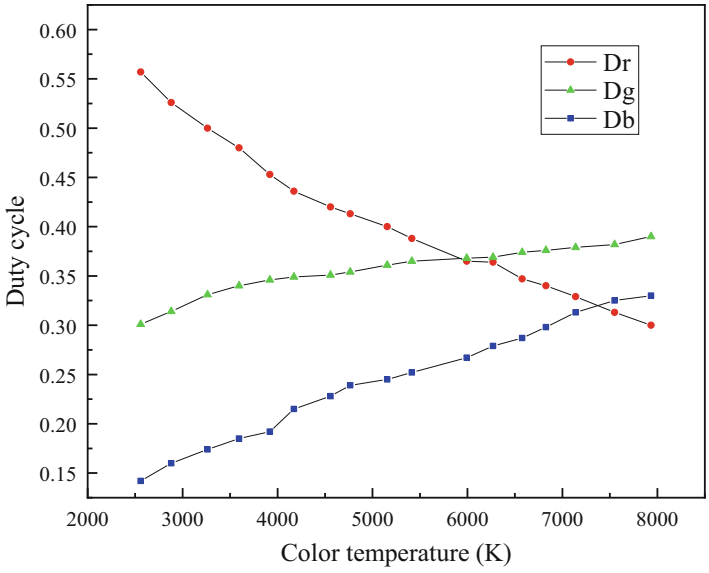


Fig. 42.10 Curve of color coordinates changing with color temperature

The irradiance range of mixed white light is 3.4–4.2 W/m^2. When the color-temperature range was 5000–6500 K, the irradiance reached its maximum and the lighting effect was optimal. The color tolerance *Duv* is stable around 0, which is specified to be less than 5.5 $\times$ 10^{-3}, and can meet the lighting standards [54]. Figure 42.13 shows that the illumination value fluctuates less with the change in color temperature, with a difference of only 8.2 lx between the maximum and minimum values. The change in color temperature has a small impact on the illumination intensity, which confirms the conclusion of the simulation.

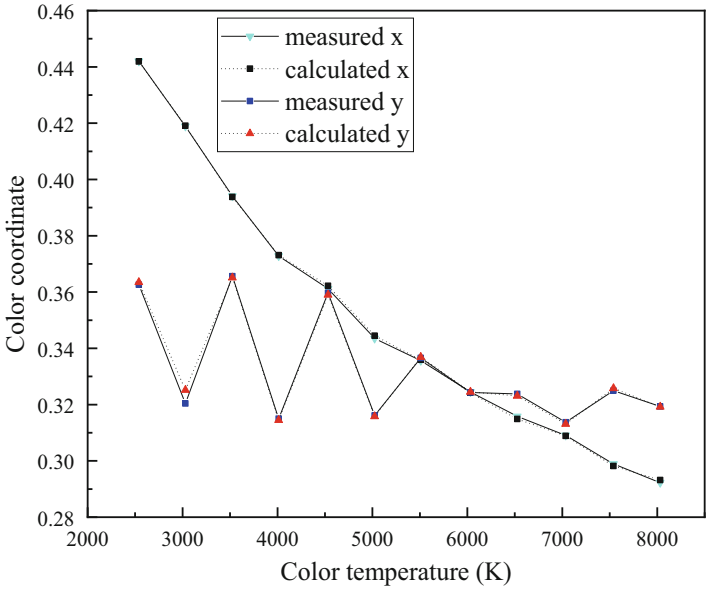


Fig. 42.11 Change of RGB duty cycle with color temperature

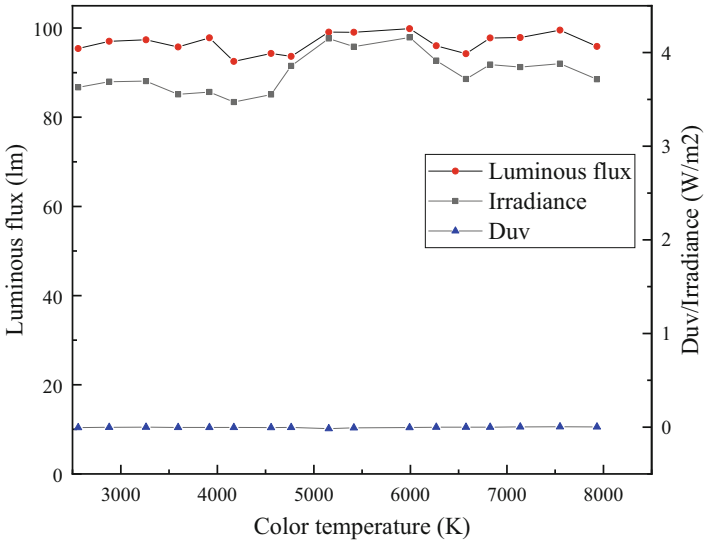


Fig. 42.12 Variation of luminous flux, irradiance, and color tolerance with color temperature

In Fig. 42.13, a comparison between the calculated color index and the measured color index according to Eqs. (42.6) and (42.7) shows that the theoretical value of the color index is in good agreement with the measured value. The color index is greatly

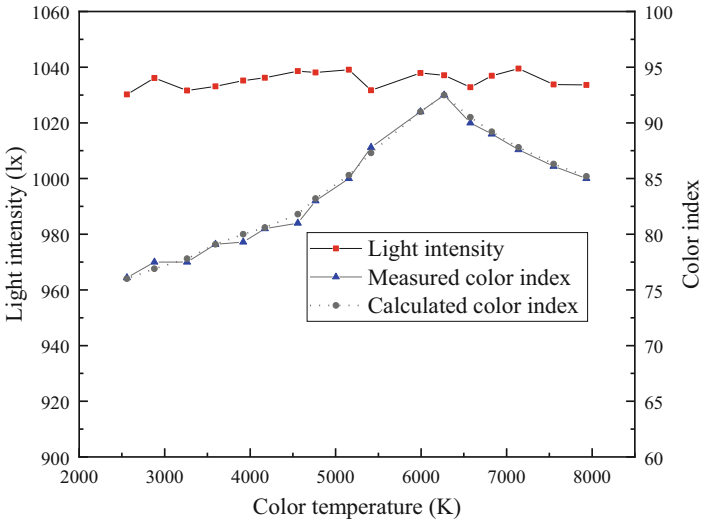


Fig. 42.13 Variation of illuminance and color index with color temperature

affected by the color temperature. When the color temperature is within the range of 5500–6500 K, the color index is optimal. This color-temperature range represents the light and color temperatures at noon. The effect of restoring the color of the object itself is the best, which also conforms to natural laws. A light source with a color index greater than 75 is of high quality [55]. The color index of the mixed white light is greater than 76 within the adjustable color-temperature range, and the maximum color index is 92, indicating that the experimental mixed white light has a good effect.

Figure 42.14 shows the light-intensity curves with distance at different color temperatures after using the HP350 spectrophotometer to measure the light intensity at different color temperatures and distances. The light-intensity curves with distance at each color temperature almost overlap, indicating that the influence of the color temperature on the light intensity is not significant. However, as the distance increases, the light intensity decreases, indicating that the distance has a significant impact on the light intensity, verifying the conclusions obtained from the simulation.

Figure 42.15 shows the light intensities at different positions on the receiving surface; they were measured at a distance of 50 cm. The light intensity is the highest under direct light. The light intensity is the lowest at the edge of the receiving surface, with a maximum illumination value of 1500 lx. The illuminance value on the receiving surface is uniform at 400–1500 lx, which meets the 300–1500 lx illuminance value formulated by the International Organization for Standardization with uniformity and lighting effects.

To verify the correctness of the simulation results, the proportions of the RGB LEDs were adjusted to 0.6: 0.28: 0.12, 0.5: 0.32: 0.19, 0.47: 0.33: 0.2, and 0.45: 0.34: 0.21, using a driving circuit, which were consistent with the simulation

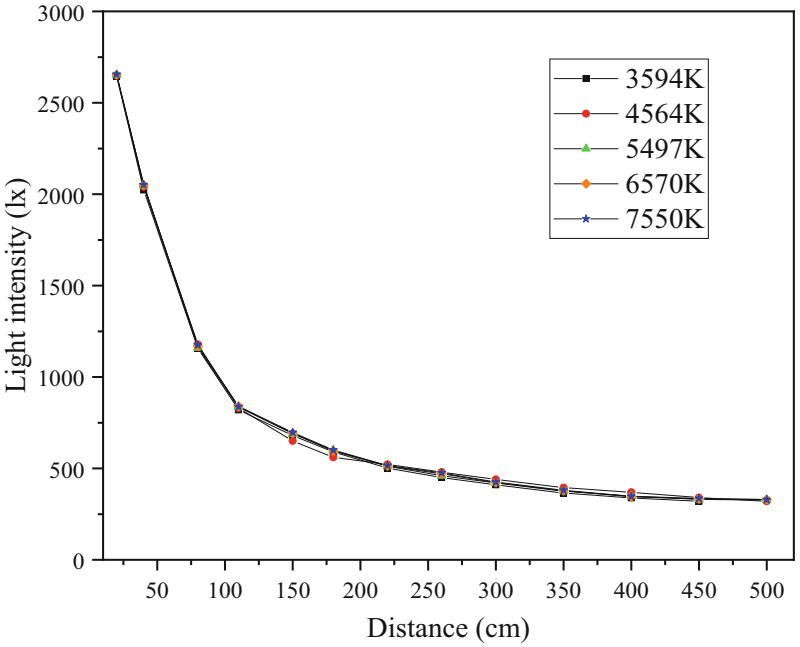


Fig. 42.14 Change of light intensity with distance under different color temperatures

proportions. The color temperatures of the mixed white light measured using the HP350 spectrophotometer were 3472, 5670, 6530, and 7260 K, which were almost equal to the simulated color temperatures. Figure 42.16 shows the measured spectrogram and Fig. 42.17 shows the position information of the mixed white light on the CIE 1931 chromaticity map. The chromaticity coordinates were (0.4043, 0.3846), (0.285, 0.3333), (0.3117, 0.3335), and (0.3038, 0.3098), which are very close to the simulated chromaticity coordinates, indicating that the mixed white light has a good effect.

Design and Experiment of an RGB-LED Visible-Light Communication System

System Working Principle and Experimental Device

The visible-light communication system consists of three parts: a sender, an indoor channel, and a receiver. Figure 42.18 shows the structural diagram of the visible-light communication system. The light emitted by the three differently colored LED chips at the sending end can be regarded as three channels; therefore, color-division multiplexing was used for multiplexing transmissions, thereby improving the system communication performance.

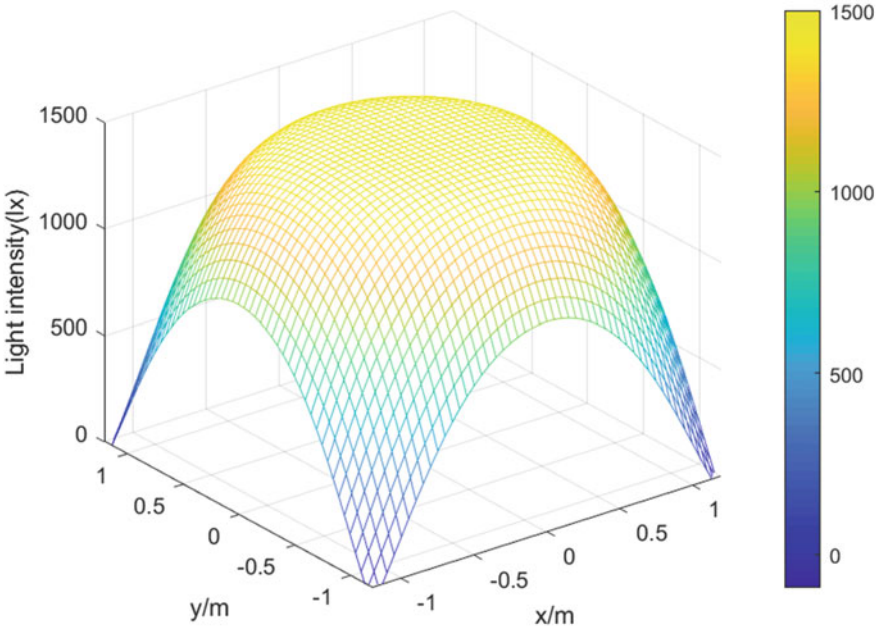


Fig. 42.15 Distribution of light intensity on the receiving surface

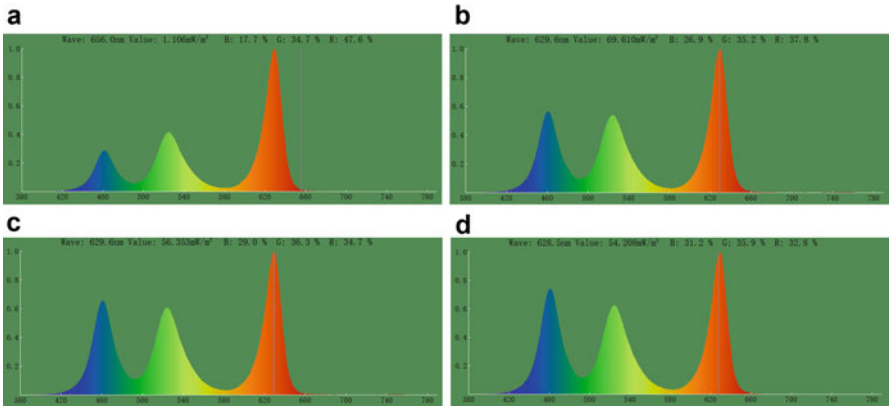


Fig. 42.16 Measured spectrograms at different color temperatures: (**a**) color temperature is 3470 K, (**b**) color temperature is 5670 K, (**c**) color temperature is 6530 K, and (**d**) color temperature is 7260 K

Three PWM channels are generated by the STM32 microcontroller to drive the three types of LEDs. The three colors are then mixed into white light in a certain proportion for output and transmission in the indoor channel. At the receiving end, monochromatic radiation is separated from the white light with corresponding

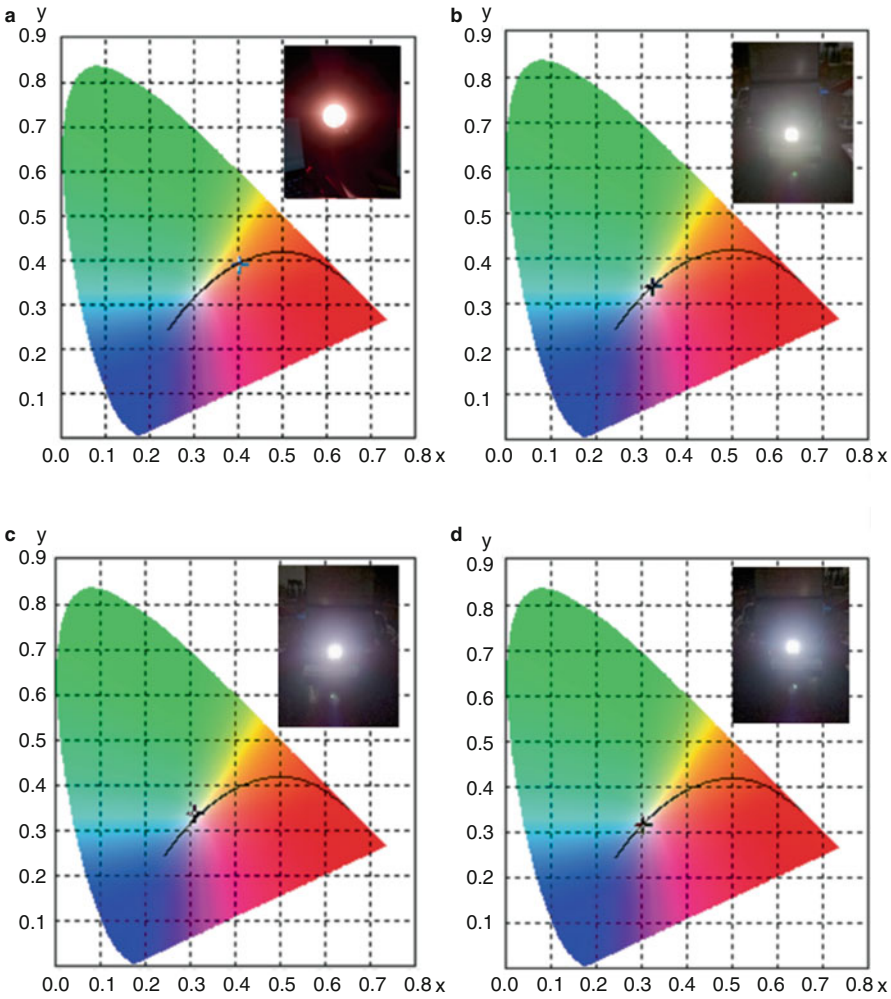


Fig. 42.17 CIE 1931 chromaticity chart at different color temperatures. (**a**) 3470 K. (**b**) 5670 K. (**c**) 6530 K. (**d**) 7260 K

filters to form three channels. The detector is used to receive signals and display them on the oscilloscope. The received optical power is measured using an optical power meter and the data are recorded using a computer. Finally, the color temperature and light intensity of the mixed white light are measured using a spectrophotometer. The RGB LED visible-light communication system designed in this study achieved high-speed data communication through three parallel signals and used the PWM dimming method to output white light to control the chromaticity.

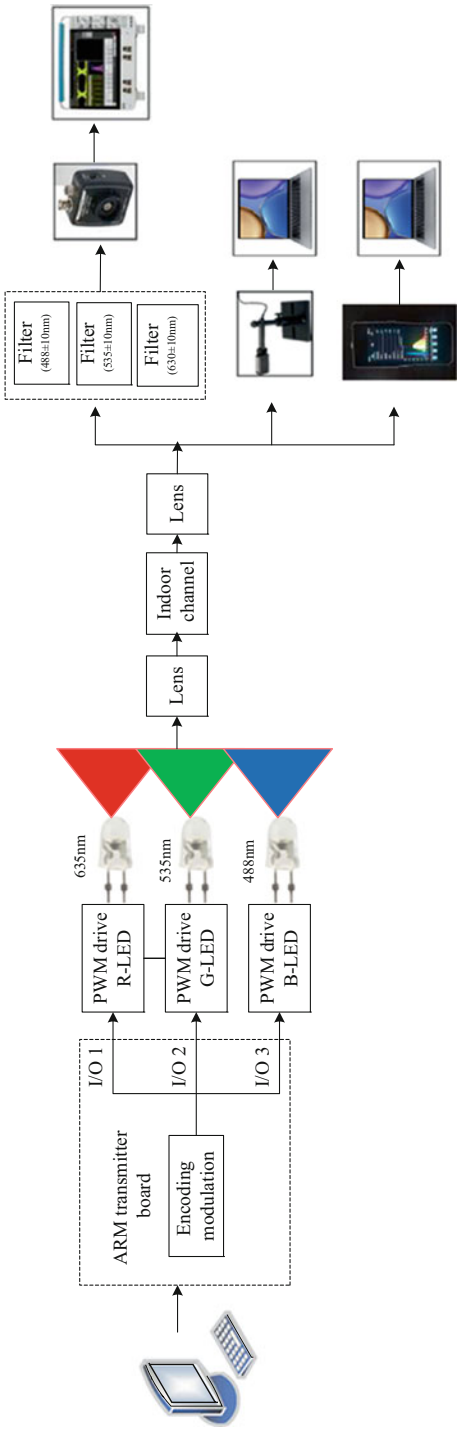


Fig. 42.18 RGB LED visible-light communication system

Analysis of Experimental Results

Because lighting and communication are integrated in this system, it is necessary to separately analyze the received light power and light intensity. The light source was vertically illuminated onto the receiving surface, and the color temperature of the light source was adjusted according to the mixing ratios obtained from the dimming experiment: 0.6: 0.28: 0.12, 0.5: 0.32: 0.19, 0.47: 0.33: 0.2, and 0.45: 0.34: 0.21. An OPHIR-PD300-UV optical power meter was used to detect changes in the light power at different distances, as shown in Fig. 42.19.

It can be observed from Fig. 42.19 that the optical power decreases with increasing distance. When the distance is less than 180 cm, the optical power rapidly decreases as the distance increases. When the distance exceeds 180 cm, the optical power decreases slowly as the distance increases. These five groups of curves generally conform to exponential-decay trends. The receiving-surface light power of the warm color temperature is relatively low, whereas the light power of the cold color temperature is higher. However, when the distance is greater than 320 cm, the light powers of the five color temperatures tend to be consistent.

Next, when the distance was 100 cm, the distribution of light power on the receiving surface under the five color temperatures was measured. The analysis shows that the distribution trend of light power on the receiving surface is consistent

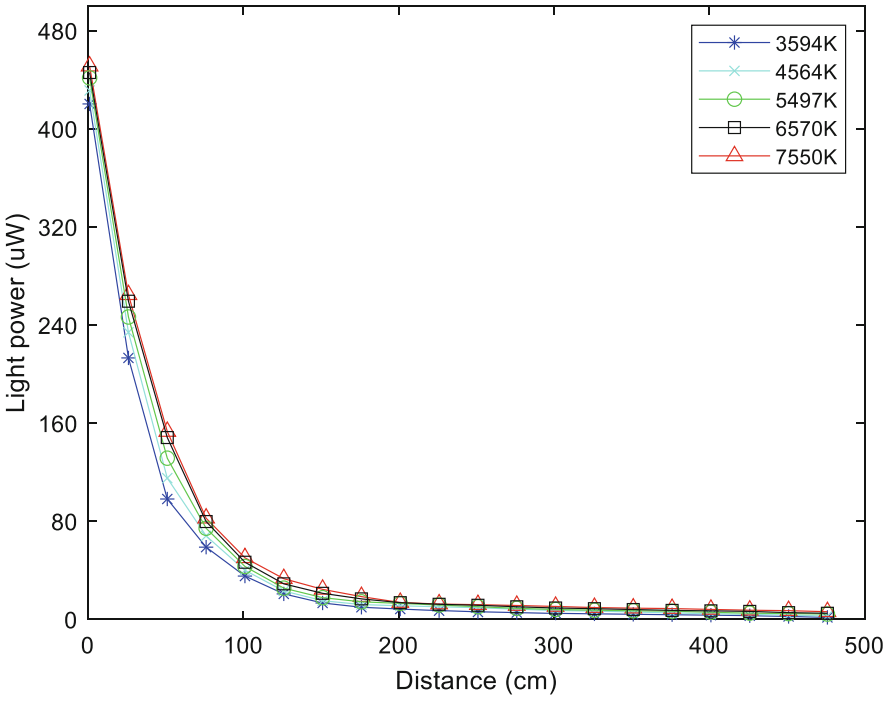


Fig. 42.19 Changes in optical power with distance

among the five color temperatures. Taking a color temperature of 6570 K as an example, as shown in Fig. 42.20, the value of the light power is the highest at the center position and decreases sequentially around the four positions.

The distance between the transmitter and receiver remained fixed, the modulation bandwidth was changed, and the error rate under different modulation bandwidths was calculated to obtain the maximum modulation bandwidth for reliable system transmissions. In the experiment, the communication distance was 250 cm, and the modulation bandwidths of the three channels were changed. The total modulation bandwidths at different color temperatures were recorded and synthesized, and the error rates were calculated.

Figure 42.21 shows the error-rate curve with the modulation bandwidth at five color temperatures: 3594 K, 4564 K, 5497 K, 6570 K, and 7550 K. As the modulation bandwidth increases, the bit-error rate also increases. When the error rate is below 10^{-3}, the maximum modulation bandwidth for the five color temperatures can reach 9.2 MHz, 10 MHz, 11 MHz, 11.5 MHz, and 11.7 MHz, respectively. The modulation bandwidth at high color temperatures is higher than that at low color temperatures. When the modulation bandwidth is 9 MHz, the error rate at all color temperatures is less than 3.8×10^{-3}, achieving reliable transmissions in visible-light communication systems.

When the modulation bandwidth is 8 MHz, the error rate curves of the five color temperatures of 3594 K, 4564 K, 5497 K, 6570 K, and 7550 K with distance are

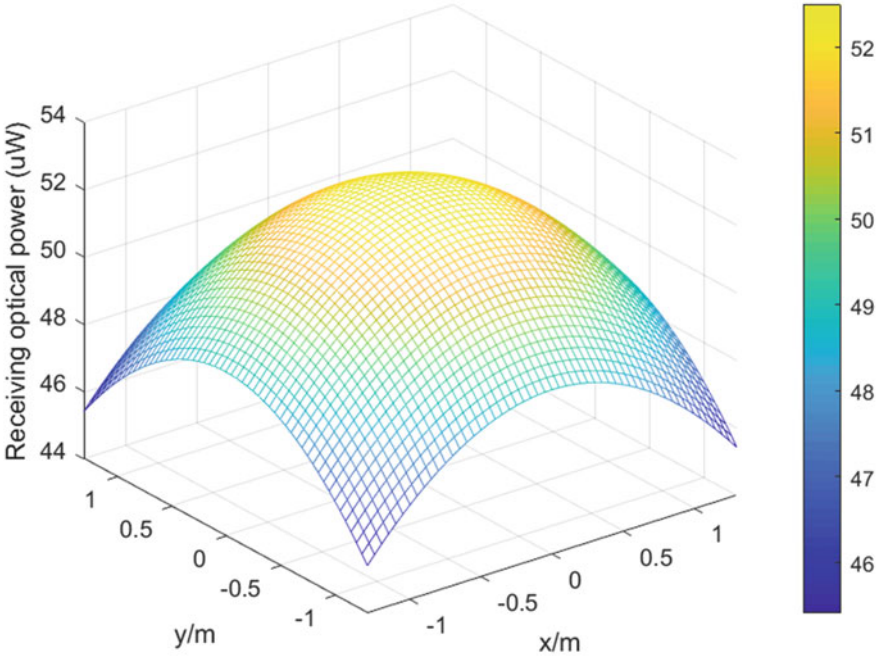


Fig. 42.20 Distribution of optical power on the receiving surface

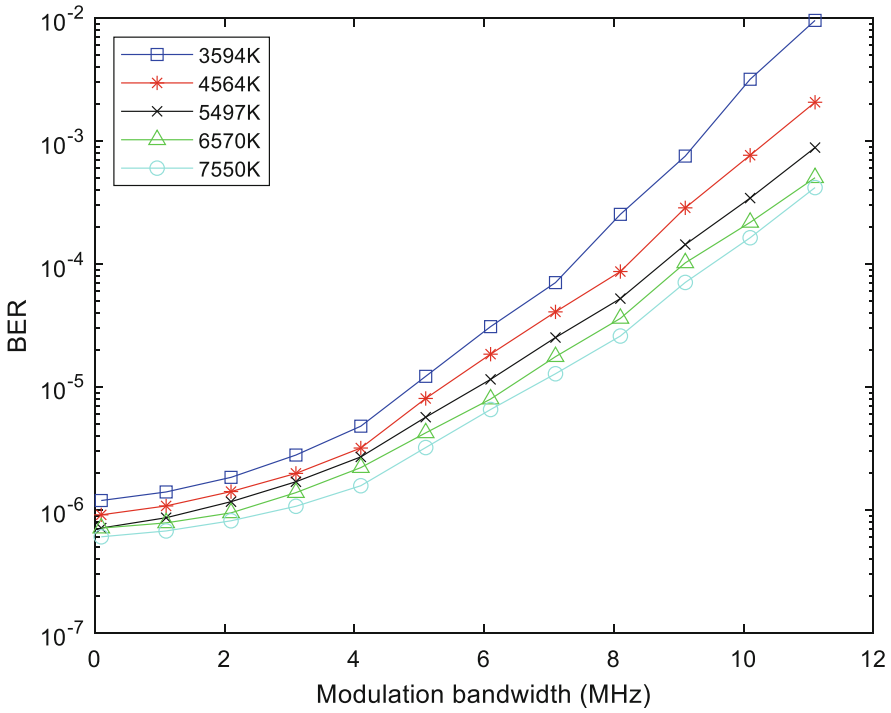


Fig. 42.21 Changes in bit-error rate with modulation bandwidth for different color temperatures

shown in Fig. 42.22. In the actual testing, as the distance increases, the error rate of the system also increases, and the speed of change increases. When the transmission distance is 350 cm, the error rate at each color temperature is lower than the limit error-rate value 3.8×10^{-3} for forward error correction. The bit-error rate of the cold white temperature is significantly lower than that of the warm white temperature.

It was concluded that as the color temperature increased, the communication performance of the system improved. However, when the color temperature exceeded 6500 K, the increase in the color temperature had a smaller impact on the communication system performance.

Conclusion

This study investigated the chromaticity control of light sources in visible-light communication systems. LED white-light sources with adjustable color temperatures were investigated through simulations and experiments. When the error rate of each color temperature was below 3.8×10^{-3}, the modulation bandwidth of the system reached 11.7 MHz, the farthest communication distance was 5 m, and the

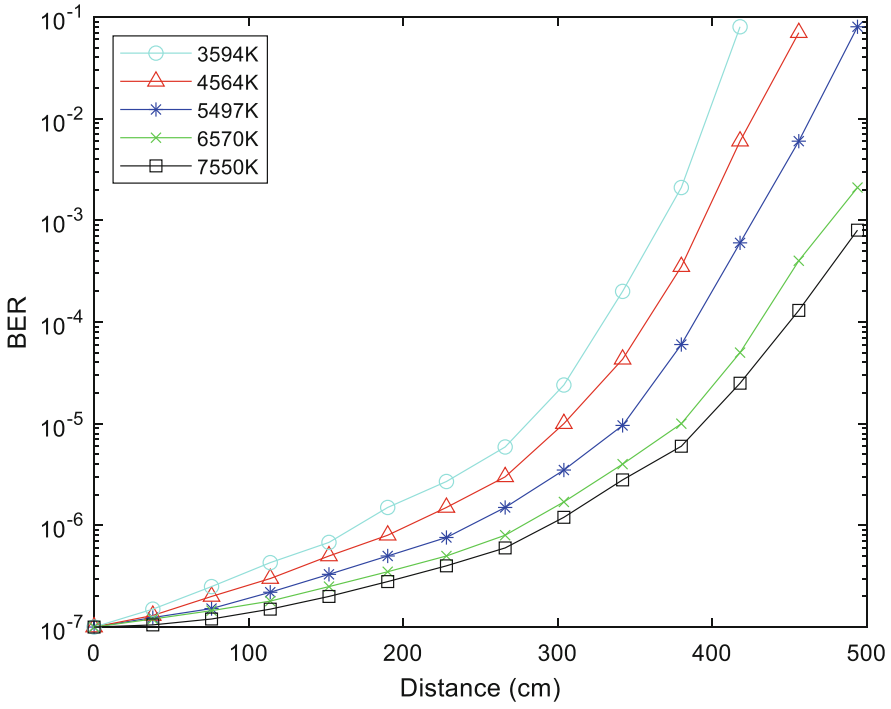


Fig. 42.22 Change of error rate with distance for different color temperatures

illumination intensity was above 400 lx. This enabled both the visible-light communication and lighting functions. The following conclusions were obtained:

1. Light sources with different color temperatures can be obtained by changing the ratio of red, green, and blue LEDs. The general color index of the light source with a color temperature between 5500 K and 6500 K was the best. The influence of the color temperature on the illumination intensity of the receiving surface at the same distance could be ignored. Distance had a significant impact on the illumination of the receiving surface. As the distance increased, the light intensity decreased rapidly.
2. The optical powers at different color temperatures and distances were measured. It was found that the color temperature had a small impact on the receiving power value, whereas the distance had a greater impact. As the distance increased, the optical power decreased exponentially.
3. Based on the communication experiments under different color temperatures, it was concluded that the communication performance of the system was better in the high-color-temperature range and significantly better than in the low-color-temperature range. As the distance increased, the communication error rates under different color temperatures continued to increase.

References

1. Favero F, Lowden A, Bresin R, et al. Study of the Effects of Daylighting and Artificial Lighting at 59° Latitude on Mental States, Behaviour and Perception [J]. Sustainability, 2023, 15(2): 1144–1165.
2. Kakitsuba Naoshi. Comfortable indoor lighting conditions for LED lights evaluated from psychological and physiological responses [J]. Applied ergonomics, 2020, 82(10): 102941–102948.
3. Mills P R, Tomkins S C, Schlangen L. The effect of high correlated colour temperature office lighting on employee wellbeing and work performance [J]. Journal of Circadian Rhythms, 2007, 5(1):1–2.
4. Cajochen C, Freyburger M, Basishvili T, et al. Effect of daylight LED on visual comfort, melatonin, mood, waking performance and sleep. Lighting Research & Technology. 2019; 51(7):1044–1062.
5. Giménez, Marina C, Geerdinck L M, Versteylen M, et al. Patient room lighting influences on sleep, appraisal and mood in hospitalized people [J]. Journal of Sleep Research, 2017, 26(2): 236–246.
6. Figueiro M G, Steverson B, Heerwagen J, et al. The impact of daytime light exposures on sleep and mood in office workers [J]. Sleep Health, 2017, 3(3):204–215.
7. Matheus L E M, Vieira A B, Vieira L F M, et al. Visible Light Communication: Concepts, Applications and Challenges [J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3204–3237.
8. Ding D Q, Ke X Z. Research on Visible Light Communication and Its Key Technologies [J]. Semiconductor Optoelectronics, 2006 (02): 114–117.
9. Ding DQ, Ke X Z. Design and Simulation of Light Source Layout for VLC System [J] Optoelectronic Engineering, 2007, 34(1): 131–134.
10. Wang Y, Wang Y, Chi N, et al. Demonstration of 575-Mb/s downlink and 225-Mb/s uplink bi-directional SCM-WDM visible light communication using RGB LED and phosphor-based LED [J]. Optics Express, 2013, 21(1):1203–1208.
11. Wu F M, Lin C T, Wei C C, et al. 3.22-Gb/s WDM visible light communication of a single RGB LED employing carrier-less amplitude and phase modulation [C]. 2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC), Anaheim, California United States, 2013: 1–3.
12. Wang Y, Tao L, Huang X, et al. 8-Gb/s RGBY LED-based WDM VLC system employing high-order CAP modulation and hybrid post equalizer [J]. IEEE photonics journal, 2015, 7(6): 1–7.
13. Huang X, Shi J, Li J, et al. A Gb/s VLC Transmission Using Hardware Pre equalization Circuit [J]. IEEE Photonics Technology Letters, 2015, 27(18):1915–1918.
14. K Y, Ke X Z. Multi-dimensional coding in visible light communication [J]. China Laser, 2015, 42(02): 138–144.
15. Zhang, Rong, Hanzo, et al. Resource Allocation Under Delay-Guarantee Constraints for Heterogeneous Visible-Light and RF Femtocell [J]. IEEE Transactions on Wireless Communications, 2015, 14(2):1020–1034.
16. N Chi, M J Zhang, Y J Zhou. 3.375-Gb/s RGB-LED based WDM visible light communication system employing PAM-8 modulation with phase shifted Manchester coding [J]. Optics express, 2016, 24(19):21663–73.
17. Chen D H, Cheng Y J, Chen Y W, et al. Phosphor-based LED Visible Light Communication System Bandwidth Enhancement Employing MC-CDMA [C]. Optical Fiber Communications Conference and Exhibition, Los Angeles, CA, USA, 2017:1–3.
18. Lu I, Lai C H, Yeh C H, et al. 6.36 Gbit/s RGB LED-based WDM MIMO Visible Light Communication System Employing OFDM Modulation [C]. Optical Fiber Communications Conference and Exhibition, Los Angeles, CA, USA, 2017:39–42.
19. Kong M, Lv W, Ali T, et al. 10-m 951-Gb/s RGB laser diodes-based WDM underwater wireless optical communication [J]. Optics Express, 2017, 25(17): 20829–20834.

20. Zhu X, Wang F, Shi M, et al. 10.72 Gb/s visible light communication system based on single packaged RGBYC LED utilizing QAM-DMT modulation with hardware pre-equalization [C]. Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, USA, 2018: 3–6.
21. Deng K, Wan Y, Lu Y. MPPM based dimming control scheme in visible light communication systems [J]. Optics Communications, 2019, 451(8): 168–173.
22. Zuo Y, Zhang J. A Novel Coding Based Dimming Scheme with Constant Transmission Efficiency in VLC Systems [J]. Applied Sciences, 2019, 9(4):803–810.
23. Hu F, Li G, Zou P, et al. 20.09-Gbit/s underwater WDM-VLC transmission based on a single Si/GaAs-substrate multichromatic LED array chip [C]. 2020 Optical fiber communications conference and exhibition (OFC), San Diego, CA, USA, 2020: 4–7.
24. Zhang Z X. Design of Wireless Visible Light Communication Control System Based on LED Lighting [D]. Harbin: Harbin Institute of Technology, 2022.
25. Qiu P, Zhu S, Jin Z, et al. Beyond 25 Gbps optical wireless communication using wavelength division multiplexed LEDs and micro-LEDs [J]. Optics letters, 2022, 47(2):317–320.
26. Min H H. Research and Implementation of Four Primary Color Silicon Based LED Color Modulation Technology for Lighting and Communication Constraints [D]. Nanchang: Nanchang University, 2023.
27. Chen T, Liu L, Hu W W. Research on Visible Light Communication [J]. ZTE Communication Technology, 2013, 19(1): 49–52.
28. Vucic J, Kottke C, Nerreter S, et al. 513 Mbit/s Visible Light Communications Link Based on DMT-Modulation of a White LED [J]. Journal of Lightwave Technology, 2010, 28(24): 3512–3518.
29. Vucic J, Kottke C, Habel K, et al. 803 Mbit/s visible light WDM link based on DMT modulation of a single RGB LED luminary [C]. Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, Los Angeles, CA, USA, 2011:1–3.
30. Kottke C, Hilt J, Kai H, et al. 1.25 Gbit/s visible light WDM link based on DMT modulation of a single RGB LED luminary [C]. 2012 38th European Conference and Exhibition on Optical Communications, Amsterdam, Netherlands, 2014. 3–4.
31. Cossu G, Khalid A M, Choudhury P, et al. 3.4 Gbit/s visible optical wireless transmission based on RGB LED [J]. Optics Express, 2012, 20(26): B501–B506.
32. Cossu G, Wajahat A, Corsini R, et al. 5.6 Gbit/s downlink and 1.5 Gbit/s uplink optical wireless transmission at indoor distances ($\geq$ 1.5 m) [C]. 2014 The European Conference on Optical Communication (ECOC), Cannes, France, 2014:1–3.
33. Cossu G, Ali W, Corsini R, et al. Gigabit-class optical wireless communication system at indoor distances (1.5—4 m)[J]. Optics Express, 2015, 23(12):15700–15705.
34. Sung J Y, Chow C W, Yeh C H. Dimming-discrete-multi-tone (DMT) for simultaneous color control and high speed visible light communication [J]. Optics express, 2014, 22(7): 7538–7543.
35. Lee S, Ahn B G, Ju M C, et al. A modified VPPM algorithm of VLC systems suitable for fast dimming environment [J]. Optics Communications, 2016, 365: 43–48.
36. Sewaiwar A, Chung Y H. High-performance TDM-MIMO-VLC Using RGB LEDs in Indoor Multiuser Environments [J]. Current Optics and Photonics, 2017, 1(4):289–294.
37. Bian R, Tavakkolnia I, Haas H. 10.2 Gb/s visible light communication with off-the-shelf LEDs [C]. 2018 European Conference on Optical Communication (ECOC), Rome, Italy, 2018: 1–3.
38. Bian R, Tavakkolnia I, Haas H. 15.73 Gb/s visible light communication with off-the-shelf LEDs [J]. Journal of Lightwave Technology, 2019, 37(10): 2418–2424.
39. Jain S, Mitra R, Bhatia V. Multivariate-KLMS based Post-Distorter for Nonlinear RGB-LEDs for Color-Shift Keying VLC [C]. 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey, 2019:1–6.
40. Viñals R, Muñoz O, Agustín A, et al. Multi-User Precoder Designs for RGB Visible Light Communication Systems. Sensors. 2020; 20(23):6836–6842.

41. Peng S F, Han Q H, Li L H, et al The effect of LED light source on the secretion of monocyte chemokine-1 and interleukin-8 by human retinal pigment epithelial cells [J] New Advances in Ophthalmology, 2016, 36(3): 201–205.
42. Wei SJ, Nan L, Wang X, et al. Study on the effect of danshensu on LED induced light damage in human lens epithelial cells [J]. Chinese Journal of Practical Ophthalmology, 2015, 33(8): 937–942.
43. Ohno Y. Color rendering and luminous efficacy of white LED spectra [J]. Optics letters, 2004, 5530(8): 88–98.
44. Zhang X, Wang Q, Li J, et al. Estimating spectral reflectance from camera responses based on CIE XYZ tristimulus values under multi-illuminants [J]. Color Research & Application, 2017, 42(1): 68–77.
45. Hu Y B, Zhuang Q R, Liu S W, et al. Research on High Color rendering Index LED Synthetic White Light Source [J]. Journal of Optics, 2016, 36(03): 216–225.
46. Chen J B, Yu J H, Gao Y F, et al. Research on LED white light sources with ultra-high color rendering index and adjustable color temperature [J]. Journal of Optics, 2015, 35(10): 252–258.
47. Xia L. Design of Intelligent LED Lighting System [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2014
48. Li Y M, Tian D Z, Wu H, et al. Factors affecting the color parameters of LED lighting [J]. Light Source and Lighting, 2018 (03): 18–22.
49. Xiong C Y, Wu Y X, Li Y, et al. Application of RGBW four color LED mixing optimization in solar simulation [J]. Acta Photonica Sinica, 2017, 46(08): 43–52.
50. Colaco A M, Kurian C P, Kini S G, et al. Thermal characterization of multicolor LED luminaire [J]. Microelectronics Reliability, 2017, 78(02): 379–388.
51. Xu D S, Chen X, Zhu X, et al. Adjustable color temperature adjustable lighting source based on cold and warm white LED [J]. Journal of Optics, 2014, 34(01): 226–232.
52. Song P C, Wen S S, Chen Y C. Research on Mixing Light Based on RGBW Four Color LED [J]. Journal of Optics, 2015, 35(09): 313–321.
53. Song P C, Wen S S, Shang J, et al. Dimming and color adjustment method for three primary color LED based on PWM [J]. Journal of Optics, 2015, 35(02): 293–300.
54. Xu X, Chen G, Gao Y. Research on Four Primary Color LED Hybrid White Light with Brightness Continuous Adjustable under Different Color Temperature [J]. Imaging Science and Photochemistry, 2018, 36(6):507–513.
55. Tian H, Liu J, Zhen H, et al. Dimming Method for R/G/B/WW Light Emitting Diode Based on Four Channels' Pulse Width Modulation [J]. Acta Optica Sinica, 2018, 38(4): 0423001–0423002.

Visible-Light Channel-Equalization Algorithm Based on the Fusion of a Neural Network and a Temporal-Feature Memory Structure

43

Contents

Introduction 1584
Research Progress 1587
Status of Traditional Equilibrium-Technology Research 1587
Balanced-Technology Research Status of Deep Learning 1587
Channel Balancing in Visible-Light Communication Systems 1590
Time-Domain Balance-Filtering Principle 1590
Communication-System Analysis and Simulation 1592
Performance Analysis of a VLC System in a Multipath Channel 1594
CNN-LSTM Neural Network Equalization Algorithm 1595
Overall Structure 1595
Robustness Analysis of the Algorithm 1598
Simulation Analysis of the CNN-LSTM Model Performance 1599
Structural Parameters of the CNN-LSTM Algorithm 1599
CNN-LSTM Model Parameter Settings 1599
Simulation Analysis of the Structural Effectiveness of the CNN-LSTM Algorithm 1600
Comparative Simulation Analyses of Different Equalization Algorithms 1606
Simulation Analysis of the Structural Parameters of the CNN-LSTM Algorithm 1608
Conclusion 1612
References 1615

Abstract

Light communication channels sometimes have changing characteristics that are difficult to predict. This article proposes a structure based on a convolutional neural network (CNN) structure, with a balanced algorithm for a residual structure. The optical-channel balance-coefficient vector is a time sequence that reflects the channel noise characteristics and has memory characteristics. The balancer algorithm can accurately learn the complex channel characteristics and calculate the compensation coefficient, based on the channel characteristics, to restore the original transmission signal. In addition, this article discusses a method of compensating for the receiving terminal when in the mobile state.

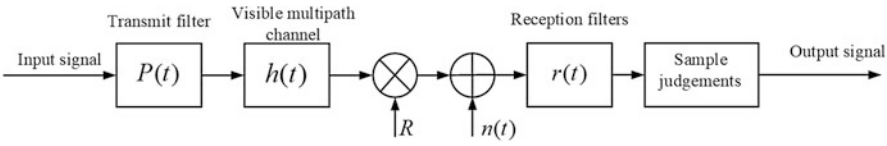
© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_43

Long short-term memory (LSTM) parameters are used to represent the sequence causal relationship in the memory channel, and the CNN and residual structure are used to refine the results to improve the reconstruction accuracy. The simulation results show that this algorithm can effectively eliminate the declining characteristics of the optical channel, improve the performance of the system-transmission error rate, solve the overall channel-damage problem, and accurately restore the original transmission signal at a faster convergence speed. In addition, compared with traditional competition-balance methods, this method achieves a better balance between performance and complexity, proving its potential and effectiveness.

Keywords

Visible light communication · Channel equalization · Residual structure · Timing feature memory structure · Convolutional neural network



Introduction

Humans are now entering the era of big data. The emergence of the concept of artificial intelligence and the popularity of mobile electronics have led to an increasingly crowded spectrum of resources in wireless communication networks. Visible-light communication (VLC) technology is fast, safe, and environmentally friendly, with no need for spectrum permits; it has become a research hotspot in the field of wireless communication.

Visible-light communication systems are susceptible to multidiameter effects, scattering noise, thermal noise from electronic devices [1], and other external factors. The intersymbol interference (ISI) phenomenon has increased, thereby increasing the system error rate. To reduce the system error rate, the balance of the light-communication channel-equilibrium technology directly corrects the communication-code elements, and eliminates direct and indirect problems between the codes.

The channel-balance algorithm is the core of the equilibrium and determines its calculation speed and performance. Traditional channel-balance algorithms include zero equilibrium, least mean square (LMS) equilibrium, recursive least squares (RLS) equilibrium, and minimum mean-square error (MMSE) equilibrium [2].

However, these traditional channel-balance algorithms have some deficiencies:

1. Relying on prior knowledge: Traditional equilibrium methods require prior knowledge, including channel-status information and noise-statistics information, which are difficult to accurately obtain in practical applications;
2. Strict assumptions about the signal model: The traditional equilibrium method usually makes strict assumptions about the signal and noise models. In actual applications, signals and noise often have unknown complex features, which may not fully meet these assumptions;
3. Sensitivity to noise conditions: Traditional equilibrium methods are more sensitive to noise conditions;
4. Require manual feature engineering: Traditional equilibrium methods usually require manual design-feature extractors to extract appropriate input features. This depends on the experience of field experts, and the feature selection and design may be difficult.

With the development of neural-network technology, a large number of intelligent equilibrium algorithms have been developed and applied to the VLC balance. Relevant research has achieved many results.

Deep learning automatically analyzes and learns data by simulating the neural network in the human brain. It can establish a more complex model through multiple levels of neural networks and identify and abstract information through a series of hierarchical learning and abstraction characteristics. Deep learning has achieved significant success in image semantic segmentation and characterization, natural language processing, and various computer science and engineering fields. In recent years, deep learning has been used for channel equilibrium and has achieved some important results.

In 2019, Lu et al. proposed a neural-network balancer based on long short-term memory (LSTM) to reduce the linear and nonlinear damage to the VLC system based on the pulse amplitude [3]. In 2021, Lin et al. proposed a nonorthodox multiple access VLC (NOMA-VLC) marker based on convolutional neural networks (CNNs) to compensate for and recover signals [4]. In a multiple-input multiple-output (MIMO) visible-optical communication system, Zhao et al. produced an adaptive artificial neural network (ANN), which was feasible and had low space complexity [5].

In 2022, Lu et al. proposed models to compensate for VLC systems based on Gaussian-based inputs and Bayesian time-domain learning [6]. Chen et al. proposed a new type of temporal–frequency attention-based convolutional neural network (TFCNN) and used it as a post-equalizer in an underwater VLC (UVLC) carrierless amplitude and phase (CAP) modulation system to compensate for the nonlinear distortion of the visible-light signal [7].

The steps abovementioned deep-learning system used to balance the channel include the following:

First, the data are pre-processed to standardize the input signal. Second, a network model is built and deep-learning technology is used to build a channel-balance

neural-network model. Third, the model is trained and evaluated with well-trained models. Fourth, the performance indicators of balanced channels are calculated, such as the misunderstanding rate for model training, to obtain pre-training models. Finally, the training model is deployed and applied to actual applications and the channel is balanced to improve system performance.

However, the CNN in the above method is a one-dimensional CNN that can be used to learn and capture the local spatial characteristics in the time-domain signal, such as frequency changes in the frequency domain or time changes in the time domain; however, changes in the information cannot be effectively tracked. A separate LSTM is suitable for modeling information, which can capture the dependency relationships between the different time steps.

However, it is difficult to effectively extract the local space characteristics of the light signal's time-domain signal. An effective method simultaneously combines the CNN-LSTM with convolution and cycling and uses the spatial locality and time-dependent relationships in the time-domain signals. Therefore, this study adds residual structures based on the CNN-LSTM and designs a type of neural-network equalizer for the receiving end. It is an active adaptive nonlinear time-domain neural network, to balance the receiving signal and suppress the intersymbol interference between the codes.

In response to the impact of the nonlinear and time-variable characteristics of indoor visible-light communication systems, an adaptive channel-equilibrium algorithm that integrates convolutional LSTM binding residual structures is proposed in this article.

Compared with the traditional channel-equilibrium method, the CNN-LSTM proposed in this study has the following advantages, in terms of the channel balance:

1. It has a data driver, which can process a large amount of data. The complex and nonlinear mapping relationship between them does not need to be prioritized; the characteristics are those of better capture signals.
2. It has better adaptability, and can better handle complex signals and noise conditions, as well as different signal models and data distributions.
3. The algorithm generally has strong universality and can be applied to different fields and tasks, including voice, image, and natural-language processing.

Compared with the ablation model, that is, the model with a separate CNN and LSTM, CNN-LSTM combines convolutional, residual, and circulating network structures and uses the local and time-dependent relationships between spaces in the time-domain signal. Among them, the CNN extracts the spatial characteristics in the time-domain signal through local perception, and the LSTM models the timing relationship through the memory unit and more comprehensively represents the information of the time-domain signal. Using the overall algorithm, the optical communication channel is more effective and balanced.

Research Progress

Status of Traditional Equilibrium-Technology Research

Equilibrium technology was used in the radio-communication field in the early stages to eliminate code interference caused by the unsatisfactory characteristics of the channel, such as the zero-forcing (ZF) [8], least mean square (LMS) [9], recursive least squares (RLS) [10], and blind equilibrium algorithms [11]. These equilibrium technologies are also applicable in visible-light communication.

Wolf et al. [12] analyzed the effect of the ZF equilibrium for a visible-optical communication system that used non-return-to-zero switch-keying control. In [13], Xie et al. discussed the multidiameter interference problem in the visible-light communication channel and studied the compensation effect of zero-forcing equilibrium, minimum equity error equilibrium, and judgment-feedback equilibrium. In [14], Komine et al. proposed an adaptive equilibrium system of LMS-based algorithms for differences in the light range between multiple light sources. The simulation results showed that the adaptive balance system reduced the impact of shadows.

In [15], Bandara et al. used the RLS equilibrium algorithm in a judgment-feedback balancer. This method reduced the length of the training sequence while maintaining the performance at the same bit-error rate (BER), thereby increasing the data-transmission rate. In [16], Chen et al. experimentally demonstrated the use of adaptive digital prebalancing technology in indoor VLC systems. The results showed that, using the proposed adaptive-balancing technology, the modulation bandwidth of the orthogonal frequency-division multiplexing (OFDM) signal could be optimized according to the estimated signal-to-noise ratio information to maximize the data rate of the VLC system.

In [17], Jin et al. applied a blind balancing scheme to the intersymbol interference of a visible-light communication system. Simulation results showed that the method effectively improved the data-transmission bandwidth and dynamically compensated for signal distortion. In [18], Cao et al. investigated the signal distortion caused by intersymbol interference in a visible optical communication system and proposed a blind balance scheme (the constant modulus algorithm (CMA)) based on the receiving signal power. An experiment was conducted to verify the effectiveness of this plan.

Balanced-Technology Research Status of Deep Learning

Machine-learning applications have demonstrated good robustness in many areas, including but not limited to medical treatment [19], finance [20], and industry [21]. Machine learning has been widely used in VLC systems and has achieved significant results. Examples include machine-learning-based light-emitting diode (LED) nonlinear compensation technology, machine-learning-based signal-detection

technology [22], machine-learning-based signal demodulation, visible-optical communication channel estimation with deep learning [23], machine-learning-based automotive visible-light communication-channel modeling [24], and machine-learning-based interior visible-light communication positioning systems [25].

Applying machine learning to VLC channel balancing has many advantages. The most significant advantage is that machine learning has strong adaptability and intelligence. It can automatically learn and adapt to the complex changes and interference of channels and improve the performance and stability of the visible-light communication system. In addition, the machine-learning balance can be trained using a large amount of data to improve the accuracy and robustness and achieve better results in practical applications.

In [26], Xiao et al. proposed a method for changing the long neural-network blind equilibrium to compensate for channel distortion caused by intersymbol interference in water acoustic communication. Simulation analysis showed that the proposed method had a good balancing effect. In [27], for the first time, Hairy et al. realized an artificial neural network (ANN) equalizer based on a visible-optical communication system considering the code performance. The results showed that the ANN equalizer had the best effect on indoor VLC systems. Furtes et al. [28] proposed a post-equilibrium method based on an underwater visible-light communication system and compared it with the Volterra-based equilibrium method. It had a better balancing effect.

Zhang et al. [29] proposed a non-data-based K-nearest neighbor (KNN) equilibrium technique to reduce the system damage to high-end modulation signals. Experiments in quadrature amplitude modulation (QAM) and 64-QAM coherent optical transmission systems showed that this method had low complexity and used zero-redundant optical-fiber nonlinear equilibrium technology that effectively improved the system performance.

In [30], Wang et al. proposed a nonlinear phase-noise discrimination method based on support vector machines (SVMs), and introduced nuclear functions to create nonlinear decision-making boundaries. Compared to the maximum-likelihood estimation algorithm, the SVM had a lower code rate at the same transmitting power, and the method complexity was lower.

In [31], Ma et al. proposed a balanced scheme based on the K-means algorithm for the nonlinear distortion caused by the components in the visible-light communication system, and an experiment was performed using the VLC system. The results showed that this method effectively compensated for the nonlinear effects, and the error-rate was significantly improved.

Lu et al. [32] introduced a pre-disturbance solution based on the K-means algorithm to reduce the nonlinear distortion in a visible-light communication system. The experimental results showed that this method reduced the system error rate by at least 50%. In their study [33], Yu et al. proposed a signal decision-making method for the density-based spatial clustering of applications with noise (DBSCAN). This was the first time that DBSCAN had been successfully applied to a four-level pulse-amplitude modulation (PAM-4) VLC system. Through experiments, the Q-factor was increased by 3.9 dB in the PAM-4 VLC system.

In [34], Miao et al. proposed a nonlinear post-balancing scheme based on deep learning (DL) to deal with severe channel damage in a DC-biased optical (DCO)-OFDM system. The simulation results showed that the nonlinear distortion of the system could be effectively compensated, and the distorted symbols could be effectively recovered. In their study [35], Liu et al. proposed to integrate feedback neural-network and recursive neural-network (RNN) deep-learning frameworks to solve the problems of nonlinear distortions and linear symbols. Simulations showed that the DL-based detector effectively enhanced the signal detection when nonlinear and linear distortions existed.

In [36], Li et al. investigated a balance method using deep neural networks (DNNs) and LMS linear equilibrium and proved its effectiveness through experiments in a 64QAM VLC system. The results showed that a 2.4 Gbit/s data-transmission rate was successfully implemented in the 1.1-m indoor VLC system, and the mixed balancer exhibited better BER performance.

Costa et al. [37] proposed a convolutional neural network (CNN) equilibrium technology suitable for visible-optical communication OFDM systems. By evaluating the amplitude of the system's average error, code rate, and error vector, this method showed a better-balanced effect than the traditional minimum secondary method. In [38], Játiva et al. proposed a channel-equilibrium solution for VLC-based underground mining systems. The BER performance of this method was better than that of the zero-forcing balance algorithm.

The aforementioned domestic and foreign studies indicate that many equilibrium algorithms currently exist. Traditional balance methods include zero-forcing equilibrium, adaptive balancing, and blind equilibrium based on LMS and RLS. Although these methods can compensate for the signals, they have some shortcomings.

For example, zero-forcing equilibrium first requires the statistical characteristics of the signal, which are often difficult to acquire in practical applications. The convergence speed of the adaptive-equilibrium algorithm is slower, and it takes longer to reach the optimal state. Additionally, the blind balancing method must satisfy certain signal characteristics; otherwise, it may lead to a decline in performance.

With its rapid development, scholars worldwide have begun to study balanced technologies based on machine learning. Compared to traditional equilibrium methods, machine-learning balancing methods have more advantages. They can automatically learn and adapt to complex changes and signal interferences, thereby improving the signal-compensation effect in practical applications.

In addition, machine-learning balance methods can improve their accuracy and robustness through training data to achieve better results in practical applications. Therefore, machine-learning methods are commonly used for extensive research on channel equilibrium.

In summary, optical communication technology is emerging in today's communication field; however, the distortion caused by indoor visible-optical communication systems in China is a challenge. Balancing technology is important for

compensating VLC channels, and is the focus of current research and the research work in this article.

With its rapid development, deep learning has been widely used in various fields, including visible-light communication, and achieved good results. Moreover, applying deep-learning algorithms to VLC equilibrium also shows good robustness. Therefore, using deep-learning methods for balancing VLC systems has important research significance.

Channel Balancing in Visible-Light Communication Systems

Time-Domain Balance-Filtering Principle

In a VLC system, the sending LED driver circuit will prepare the LED to send the information, and it is launched in the form of a light signal. This modulation method can modify on–off keying (OOK) in light communication; that is, the light signal being on or off indicates binary information. "1" and "0" have a variety of modulation and encoding methods, which mostly use parallel communication [39].

The modulated visible-light signal reaches the receiving end after a space transmission. The photoelectric detector and signal-processing circuit of the receiving end perform photoelectric transformation, amplification, sampling, demodulation, and framework processing on the signal, and the original information sequence is finally restored [39].

The total impact response of the system is

$$h_i(t) = p(t) \otimes h(t) \otimes r(t) + n(t), \tag{43.1}$$

where $h_i(t)$ is the general channel impact response, $p(t)$ and $r(t)$ are nonlinear noises, such as the thermal noise of the sender and the receiving end of the electronic device, respectively, $h(t)$ is the noise of the multidiameter channel response, and $n(t)$ is the visible-light communication background noise.

The submissive response of the visible-light channel can be modeled as a random variable for multidiameter superpositioning. Figure 43.1 shows the normalized channel impulse response at different times, where *a*, *b*, *c*, and *d,* respectively, represent the channel characteristics at times T_0, T_1, T_2, and T_3. The corresponding parameter *d* is the straight-line distance from the light source to the acceptable end, *h* is the wall-reflection coefficient, *m* is the Langbo model: 1[$x = 1, y = 1, z = 3$], 2 [$x = 1, y = 4, z = 3$], and 3[$x = 4, y = 1, z = 3$].

Owing to the movement of the receiver, the multidiameter structure in the four subdiagrams changes; however, it can be seen from the figure that the overall

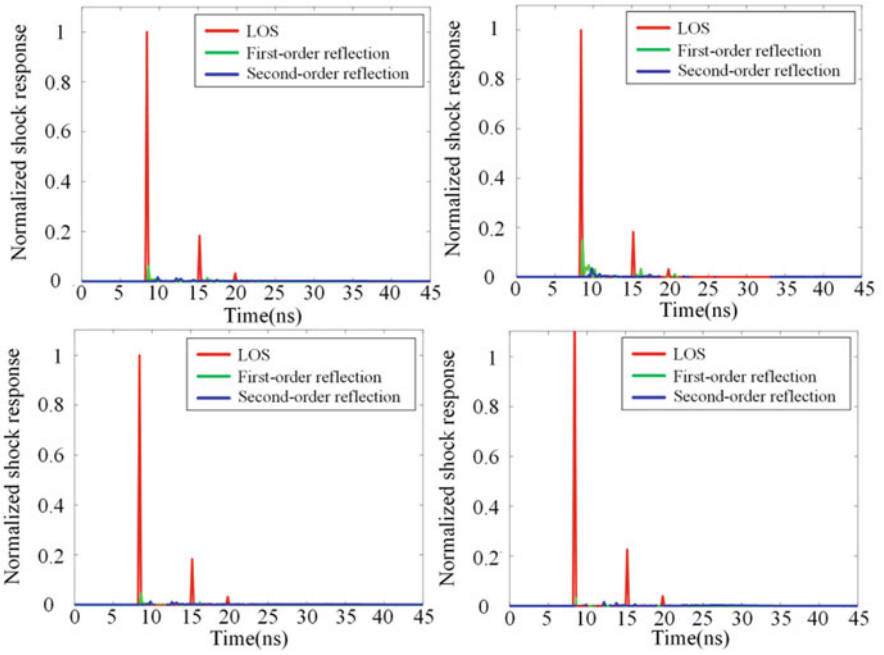


Fig. 43.1 Channel shock-response changes

structure of the visible-optical channel in different positions is relatively stable. This is because the scene established in this article is a closed indoor scene. The multidiameter component power is mainly derived from the wall reflection. Therefore, the channel is a quasi-static channel with low time-variant characteristics.

The general channel impact response of the system is distorted by various noises in the code. The principle of time-domain equilibrium is to use an adjustable-coefficient filter to adjust the impact response to satisfy the Nyquist theorem:

$$h_T(t) = \sum_{n=-\infty}^{+\infty} w_n \delta(t - nT_s), \tag{43.2}$$

where w_n is the balancing coefficient and T_s is the sampling interval. Therefore, the change in $h_i(t)$ adjusts the balancing coefficient. The tap coefficient can be determined using a filter and applied to the input signal. The filter structure is shown in Fig. 43.2.

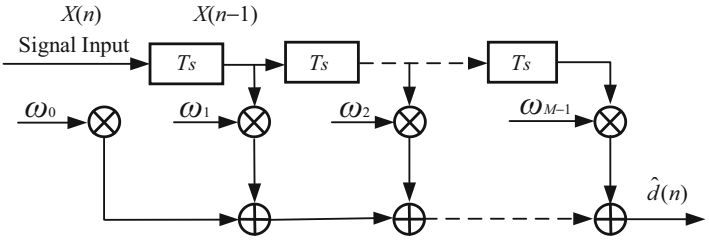


Fig. 43.2 Nonlinear horizontal time-domain filter

The input data of time n is

$$x(n) = [x(n), x(n-1), \cdots, x(n-M+1)]^T. \tag{43.3}$$

The balancer coefficient is

$$w(n) = [w_0(n), w_1(n), \cdots, w_{M-1}(n)]^T. \tag{43.4}$$

The output signal of the balancer can be represented as

$$r(n) = \sum_{n=-\infty}^{+\infty} w_i(n)x(n-i) = w^T(n)x(n). \tag{43.5}$$

The selection of the balancer coefficient directly affects the output result, which determines whether the signal can be judged correctly. In particular, if the equilibrium coefficient can adaptively change according to the channel characteristics, the balancer can dynamically correct the distorted waveform within the correction system.

Communication-System Analysis and Simulation

In a visible-light communication system, the channel capacity of the direct-current biased optical-OFDM (DCO-OFDM) transmission method is more capable than asymmetrically clipped optical OFDM (ACO-OFDM), has a doubled channel capacity, and very good application prospects. In addition, the data-transmission method of the equalizer experimental system, combined with 16QAM modulation, is used in this study. The basic principle of the DCO-OFDM system is shown in Fig. 43.3.

After the input source signal is modulated by 16QAM, a constellation mapping operation is performed. The mapping output of the DCO-OFDM system is

$$X(n) = [0, x(1), \cdots, x(N/2-1), 0, x^*(N/2-1), \cdots, x^*(1)]^T \tag{43.6}$$

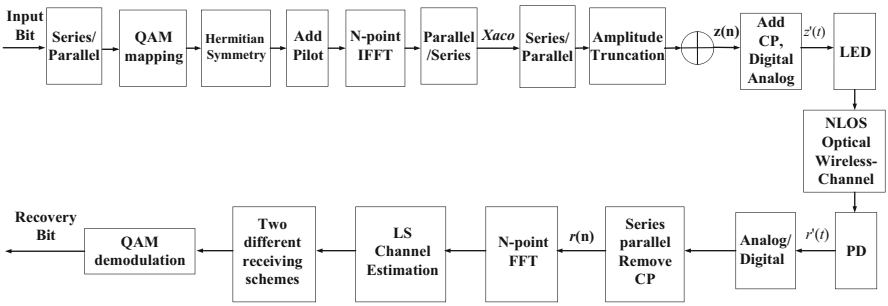


Fig. 43.3 DCO-OFDM signal-transmission principle

where N is the length of the fast Fourier transformation, * is the comortar operation, and the mapping signal meets Emi's symmetry. The next step is to use a Fourier transform on the frequency-domain mapping signal:

$$\begin{aligned} x_{\mathrm{IFFT}}(k) &= F^H X \\ &= \frac{1}{\sqrt{N}} X(0) + \frac{1}{\sqrt{N}} \sum_{n=0}^{N/2-1} X(n) \exp\left(\frac{j2\pi nk}{N}\right) + \\ &\frac{1}{\sqrt{N}} X\left(\frac{N}{2}\right) \exp(j\pi k) + \frac{1}{\sqrt{N}} \sum_{n=0}^{N/2-1} X^*(m) \exp\left(-\frac{j2\pi mk}{N}\right) \end{aligned} \tag{43.7}$$

where H represents the matrix conjugate transpose, and F is the $N \times N$ normalized discrete Fourier-transform matrix. In an intensity-modulation/direct-detection (IM/DD) system, the signal driving the LED must be a real non-negative signal; hence, the signal needs to be limited. The limited signal is converted into a parallel series and a cyclic prefix is added. After the protection interval is inserted, the digital-to-analog conversion drives the LED to emit an optical signal.

The optical signal is transmitted through the VLC indoor channel and is converted into an electrical signal by a photodetector at the receiving end. The electrical signal can be expressed as follows:

$$y(t) = h_i(t) \otimes x(t) + n(t), \tag{43.8}$$

where $h_i(t)$ is the channel impulse response proposed in section "Time-Domain Balance-Filtering Principle," $x(t)$ is the transmitted data signal, and $n(t)$ is background white-light noise. The indoor channel-noise model follows a Gaussian distribution with mean of 0.1 and variance of 1 [40]. The distribution of the Coriolis model is consistent with the background of visible-light communication in outdoor and vehicle networking.

The channel is composed of the direct and sequentially reflected channels, and its noise characteristics follow a Gaussian distribution. Because of the change in the

receiver position, the transmission distance of the channel will change. The distances of the direct channel are 2 and 2.5 m, and the distances of the reflected channel are 5 and 4.5 m, respectively. The mean values of the Gaussian distribution are 0.1 and 0.15, respectively. The variance is 0.2, 0.22.

The input signal is 182, a fast Fourier transform module and the output signal is a frequency domain.

$$Y = HX + HN_{\mathrm{DCO}}, \tag{43.9}$$

where $H = F\widetilde{h}(t)F^{H}$ is the frequency-domain expression of the channel impulse response; it is an $N \times N$ diagonal matrix. N is the Fourier transform of the additive white Gaussian noise (AWGN) vector. The 1–N/2–1 parts of the signal are extracted as the subcarrier modulation signal, and the extraction vector is

$$Y_{\mathrm{ex}} = [Y(1), Y(2), \cdots, Y(N/2-1)]^{T}. \tag{43.10}$$

The signal input to the filter is

$$Y_{\mathrm{sd_DCO}} = H_{\mathrm{sd_DCO}} X_{\mathrm{ex_DCO}} + \sigma_{\mathrm{AWNG}} H_{\mathrm{sd_DCO}} N_{\mathrm{ex_AWNG}}, \tag{43.11}$$

where $H_{\mathrm{sd_DCO}}$ is the (N/2–1) × (N/2–1) equivalent frequency-domain transfer matrix and $X_{\mathrm{ex_DCO}}$ is for X to obtain masks of $m = 1, 2, 3\ldots, N/2–1$ dimensions consisting of row elements. $N_{\mathrm{ex_AWGN}}$ is Gaussian noise with a mean of 0.1 and a variance of 1.

In the offline training phase, we used the DCO-OFDM signal-transmission mode. The model was trained in the IM/DD channel, including LED nonlinearity, optical propagation, and channel noise. First, random data were generated as transmitted symbols. After the symbols were modulated, mapped, and transmitted through the VLC channel, the received electrical signals were collected at different partial-discharge positions and divided into several short overlapping samples. Each short sequence and sample of the original transmitted signal were combined into the training data. In practical applications, it is necessary to collect a variety of training sets to improve the parameter-learning ability of the CNN-LSTM.

Performance Analysis of a VLC System in a Multipath Channel

An indoor visible-light communication system was simulated and modeled. The length, width, and height of each room were recorded as lR, bR, and hR, respectively. An LED lamp was installed at the center of the ceiling. The distance from the ceiling was marked as HLED, and its direction pointed vertically to the ground. The photodetector (PD) was located on the work plane, which was HPD from the ground, and its direction was vertically upward. The wall was divided into a grid of equal horizontal and vertical 0.1-m distances, and each cell was used as a reflection unit. The simulation parameters are listed in Table 43.1.

Table 43.1 Channel-simulation parameters

Parameter	Symbol	Numerical value
Room size	$l_R \times b_R \times h_R$/m	$6 \times 6 \times 4$
LED distance from the ceiling	H_{LED}/m	0.5
PD distance from the ground	H_{PD}/m	0.85
Monte Carlo simulation times	N_{MC}	10,000
Subcarrier number	N_{SC}	256
Cyclic prefix	N_{CP}	16
Modulation bandwidth of LED lamp	B_L/MHz	50
Photoelectric conversion coefficient	γ /(A/W)	1
PD area	A_R/cm^2	1
Area of the microreflection unit	ΔA/m^2	0.01
Wall reflectivity	ρ_i	0.8
PD field-of-view angle	Ψ_{FOV}/°	80
LED half-power angle	$\theta_{1/2}$/°	60

When the modulation bandwidth of the light source is 50 MHz, the symbolic period of the modulated signal is 10 ns. Relative to τ0, an optical signal with a delay greater than Tsp = 5 ns leads to ISI. The optical power of the receiver, line of sight (LOS), specular reflection, and other primary reflection signals exceeded 90%. Therefore, only the LOS, specular reflection, and primary reflection were considered in this experiment. In addition, considering the symmetry of the multipath effect in the room, the multipath effect of the PD on one quarter of the room was considered.

CNN-LSTM Neural Network Equalization Algorithm

Overall Structure

The convolutional neural-network–time-series memory-structure equalization algorithm can simultaneously compensate and equalize code errors caused by intersymbol interference and nonlinear noise. The basic algorithm structure is shown in Fig. 43.4. The CNN neural-network structure has a strong data-feature extraction ability; hence, it can effectively extract channel features and calculate and generate the tap coefficient $w(n)$. The process of generating the tap coefficient is also the process of classifying the 0 and 1 characters [41].

First, the channel-transmission symbols are sampled to obtain the original sequence $[X(n-(l+1))]$. As the block pilot is added during the sampling process, the sampling sequence is divided into a series of subsequences of length $2l+1$, and a two-dimensional sequence matrix is formed by the subsequences. The matrix is sent to the convolutional neural network to learn the channel characteristics and estimate the equalization coefficients [41]:

$$Z_i = w \cdot X + b, \tag{43.12}$$

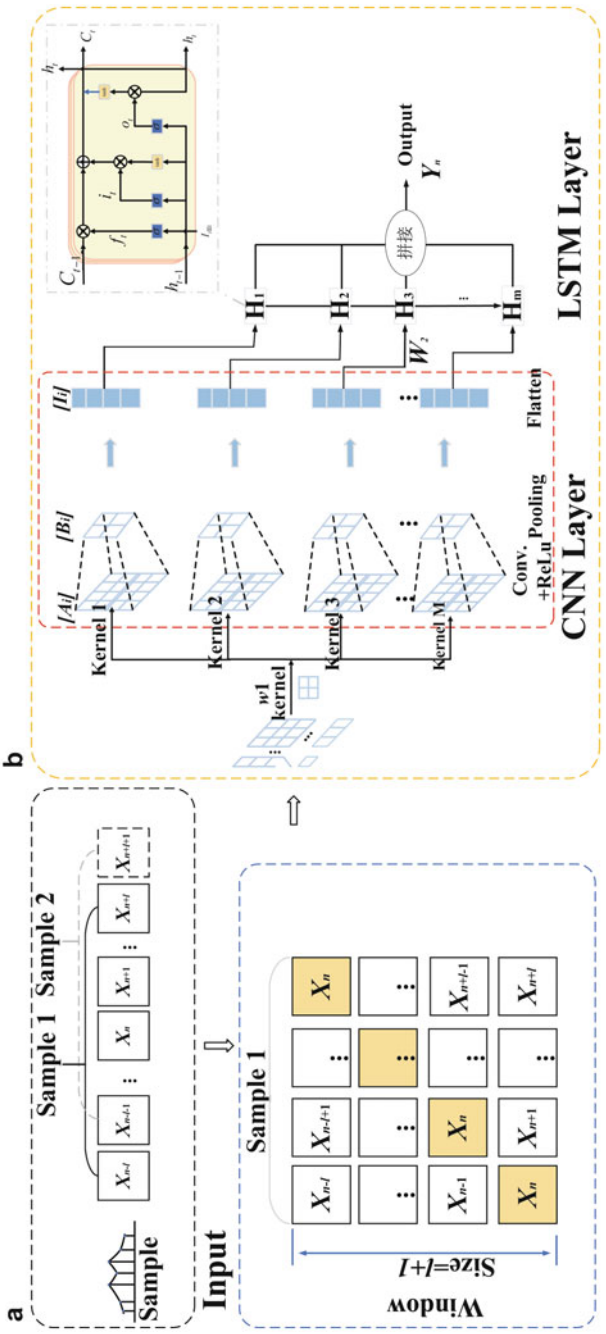


Fig. 43.4 CNN-LSTM channel equalization principle. (**a**) Pilot sampling process. (**b**) CNN-LSTM structure

where matrix X is the parameter-training matrix and b is the parameter offset vector. In addition, each neural-network layer determines the final predicted value according to the cost function. The selected cost function is the cross-entropy function:

$$L_i(w_h, w_r) = -\sum_{i}^{m} \left[w_r^i * \log\left(w_h^i\right) + \left(1 - w_r^i\right) * \log\left(1 - w_h^i\right) \right], \tag{43.13}$$

where L represents the training layer, w_h represents the predicted value, w_r represents the real sample value, and the sample w and bias b are random numbers at the initialization. Using the random gradient-descent algorithm, w and b were obtained using the minimum cost function.

In the random gradient-descent backpropagation process, the derivative of the loss function is gradually moved forward by the output layer [41]:

$$\frac{\partial L}{\partial w_h} = \left(-\frac{w_r}{w_h} + \frac{1 - w_r}{1 - w_h} \right) \cdot \frac{\partial w_h}{\partial Z^L}. \tag{43.14}$$

The existing predicted values were modified and updated according to the backpropagation and derivation results. The result of the output sequence entered the pooling layer, which sampled the result again, after reducing the sampling frequency and obtained the sampling result $[Z'_i]$. The flattening layer turns the two-dimensional matrix-sequence data into a one-dimensional vector $w_i(n)$, and each vector is input into the corresponding LSTM unit structure.

$$w_i(n) = [w_i(1), w_i(2), \cdots, w_i(n)]^T \tag{43.15}$$

The LSTM introduces a "gate mechanism" to extract the long-term memory time-series information. By controlling the input, forgetting, and output gates, the LSTM can add or remove present or past time-state information to or from the storage unit. The LSTM structure is composed of a forgetting gate, input gate, and output gate. The forgetting and input gates control the working state of the memory. The corresponding expressions are [42]

$$f_t = \sigma\left(W_f[x_t, h_{t-1}] + b_f\right) \tag{43.16}$$

$$i_t = \sigma(W_i[x_t, h_{t-1}] + b_i) \tag{43.17}$$

$$C_t = f_t C_{t-1} + i_t \tanh(W_c[x_t, h_{t-1}] + b_c), \tag{43.18}$$

where σ is an S-shaped function; W_f, W_i, and W_c represent the input weights corresponding to different doors; h_{t-1} represents the output of the previous state; and x_t represents the input of the current state. b_f, b_i, and b_c correspond to the bias matrices. The final output of the LSTM structure is

$$h_t = \tanh(C_t)(W_o[x_t, h_{t-1}] + b_o). \tag{43.19}$$

The output of the entire LSTM structure was finally activated using the rectified linear-unit (ReLU) function. After the backpropagation algorithm completed the training and optimization of the dataset, the equalizer obtained the equalization coefficient for the output-symbol equalization from the LSTM structure. It does not need to be constantly trained through the neural network, which greatly reduces the running time of the program.

When the equalization coefficient changes, owing to a change in the channel characteristics, the cost function of the LSTM structure activates the forgetting gate and neural network to re-learn the channel characteristics. The equalizer can independently adjust the equalization coefficient, according to the changes in the channel, for adaptive equalization. Based on the overall algorithm, a residual module was added to more effectively equalize the visible-light communication channel.

Robustness Analysis of the Algorithm

Visible-light communication channels vary over time. Based on the nonlinear composite channel-equalization principle of CNN-LSTM, an intelligent residual channel-equalization algorithm was proposed that could adaptively equalize the channel. However, when sampling the channel transmission symbols to obtain the original sequence $[X(n–(l+1))]$, there is often noise, and then there is a point in time when the actual input data of n is:

$$\begin{aligned} x(n) = {} & [x(n), x(n-1), \cdots, x(n-M+1)]^T \\ & + [\varepsilon(n), \varepsilon(n-1), \cdots, \varepsilon(n-M+1)]^T. \end{aligned} \tag{43.20}$$

The equalizer coefficient is

$$w(n) = [w_0(n), w_1(n), \cdots, w_{M-1}(n)]^T. \tag{43.21}$$

The output signal of the equalizer can be expressed as

$$r(n) = \sum_{n=-\infty}^{+\infty} w_i(n)(x(n-i) + \varepsilon(n-i)) = w^T(n)x(n). \tag{43.22}$$

From the above formulas, it can be observed that the selection of the equalizer coefficients affects whether the signal can be judged correctly. The original sequence sample will introduce all types of noise. The CNN has a self-learning ability to adapt to the generated noise; it balances the channel noise of different interferences in the data-feature extraction, and finally outputs a set of equalizer coefficients.

The LSTM structure further processes this group of equalizer coefficients to obtain ones that match the current channel, which improves the bit-error rate of the system transmission. In addition, a residual module is added so that the LSTM structure can better activate the CNN network structure to re-learn the group of equalizer coefficients, according to whether the input of the current equalization

coefficient matches the channel, and iterate the input to obtain the adaptive equalizer coefficients.

This design has good algorithm robustness. The normal channel-noise CNN self-learning structure can output a set of effective equalizer coefficients and the LSTM can directly handle the output. When there is abnormal noise interference, the LSTM structure will trigger the CNN self-learning structure to train again.

Simulation Analysis of the CNN-LSTM Model Performance

Structural Parameters of the CNN-LSTM Algorithm

In view of the characteristics of the time-domain signal of visible-light communication and the problems of the model, this paper proposes a deep-learning channel-equalization prediction model based on a convolutional long short-term memory network for feature fusion; it is called the CNN-LSTM equalizer for VLC channel equalization.

The channel-equalization calculation process of the CNN-LSTM equalizer is as follows: The input samples are extracted from the spatial domain by the CNN module, the resulting features are input into the LSTM for time-series feature screening, and finally, the regression layer is used for regression.

At present, there is no recognized optimal deep-learning model structure for channel equalization in visible-light communication, and different structures have a great impact on the results; therefore, choosing a good structure strongly affects the channel-equalization performance.

The LeNet-5 two-dimensional convolution neural network has been used in some fields, including image processing, and achieved good results. Therefore, this study designed the convolution layer of the CNN based on the LeNet-5 structure. It replaced the fully connected layer in the CNN with the LSTM layer, and introduced batch normalization, a pooling layer, and a discarding layer from reference [43]. The model structure is shown in Fig. 43.5.

The CNN-LSTM-equalizer model uses the signal after passing through the channel and optical medium, and outputs the received signal obtained after equalization before making a decision.

CNN-LSTM Model Parameter Settings

In this section, the performance of the proposed scheme is evaluated, based on simulation results. Modulated DCO-OFDM symbols containing 512 subcarriers with 16 quadrature amplitude modulation (16QAM) are randomly generated at the transmitter, and the normalized DC is set to 0.5. The sampling symbol is divided into a two-dimensional matrix using a block pilot, completing the data preprocessing.

The CNN-LSTM-equalizer structure consists of four CNN-LSTM-series structures. One pooling layer, one convolution layer, and one flattening layer are used in the CNN structure.

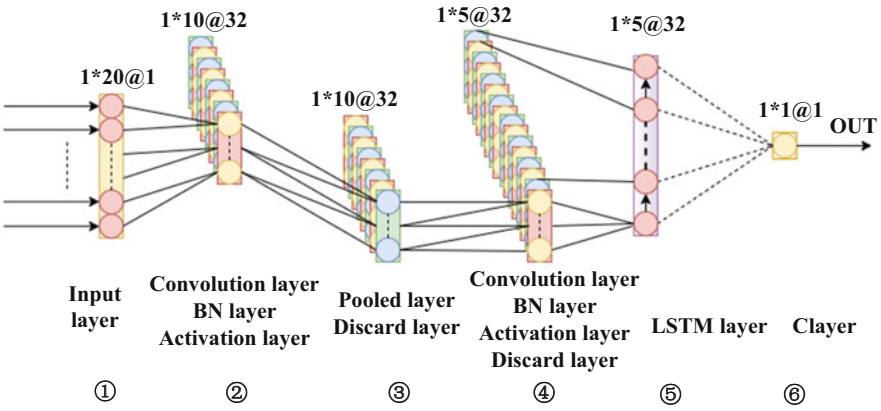


Fig. 43.5 CNN-LSTM structure diagram

Table 43.2 Equalizer simulation parameters

Parameter	Value
Modulation method	DCO-OFDM/QAM
Channel capacity	512 × 16
DC	0.5
Number of neurons in the pooling layer	32
Number of neurons in the convolutional layer	32

The numbers of neurons corresponding to the pooling layer and convolution layer are set to 32. The equalizer simulation parameters are listed in Table 43.2.

To compare all the models as closely as possible, they are run in an environment with the same training-process configuration, referring to the parameters commonly used in deep-learning model training [44]. The Kaiming method is used to initialize the weight coefficients of the CNN and full-connection layers [45]. The orthogonal method is used to initialize the weight coefficients of the LSTM layer [46], which is helpful for accelerating the model convergence.

In the network training process, the Adam optimizer makes the model converge faster and more robustly [47] and reduces the test of the verification set. The processed dataset was randomly scrambled and divided into a training set, verification set, and test set at 10:1:2, as shown in Table 43.3.

Simulation Analysis of the Structural Effectiveness of the CNN-LSTM Algorithm

Using the equalizer parameters listed in Table 43.1, the performance of the algorithm-only equalizer proposed in this study was simulated when the channel distance was 3.5 m, the normalized LED power was 1, and the power distribution

Table 43.3 LSTM, CNN, and CNN-LSTM model training-parameter configurations

CNN-layer initialization method	LSTM-layer initialization method	Full-connection-layer initialization method
Kaiming	Orthogonal	Kaiming
Optimizer method	Learning rate	Loss function
Adam	1e-3	Mean-squared error (MSE)
Number of training sets	Number of validation/test sets	Training wheels
1e4	1e3/1e3	100
Discard rate	Lot size	Activation function
0.4	32	ReLU

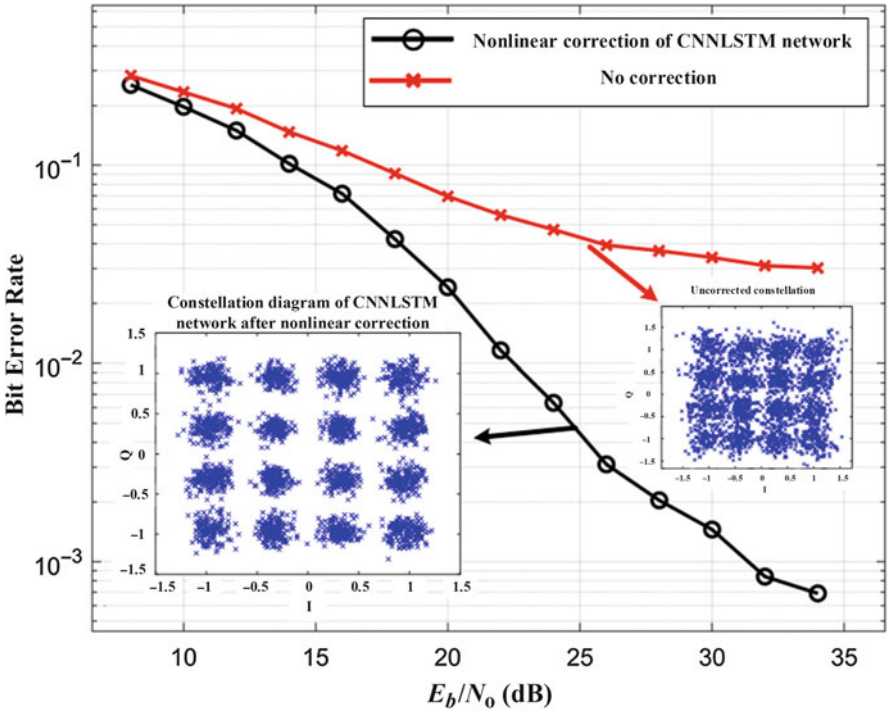


Fig. 43.6 Performance comparison of the CNN-LSTM network with and without correction

was 0.5. The simulation results are shown in Fig. 43.6. The BER change under a certain signal-to-noise ratio indicates the equalization performance [48].

In addition, this paper describes the output of the proposed scheme and ideal equalization for comparison, which reflects the channel characteristics, and compares the effect after equalization by the actual reception amplitude, ideal reception amplitude, and reception amplitude after equalization. The original Rx sequence and the sequence after equalization are shown in Fig. 43.7.

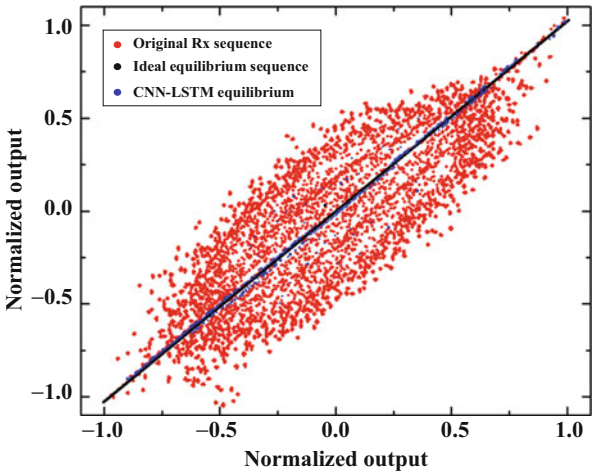


Fig. 43.7 Comparison of the normalization degree between the input and output signals

After mapping the amplitudes of x(n) and y(n), the signal-mapping amplitudes are shown in Fig. 43.7. Ideally, the receiver can fully measure the signal channel characteristics. We can observe that the originally received amplitude n deviates from a linear straight line and exhibits strong distortion. However, the proposed scheme differs slightly from the ideal case and shows excellent linearity for the same channel input. This indicates that the proposed CNN-LSTM can quasi-match the original input, thereby mitigating the general nonlinearity of the LED and optical radio channels.

Figure 43.8 shows the eye diagram of the VLC system with a 20 Mbit/s, 16QAM input signal with and without equalization. It can be seen that the eye diagram is almost closed before equalization, and the jitter and noise are relatively high, indicating that the BER is relatively high at this time. After equalization, the open eyes can be clearly observed, which indicates that the signal after equalization receives less interference from other sources, and the signal quality is higher.

Mobile Receiver-Compensation Method Based on the CNN-LSTM Algorithm Structure

In a visible-light indoor communication system, it is common for the receiver PD to receive a signal in the mobile state. Therefore, this study fully considers this scenario and aims to improve the understanding of the actual communication environment through an in-depth study to enable better design and optimization of the communication system.

By analyzing and exploring this scenario, we can encounter many types of the complex situations found in an actual communication environment, such as channel noise and multipath effects, to understand them more comprehensively. We then propose effective solutions to these problems to improve communication-system

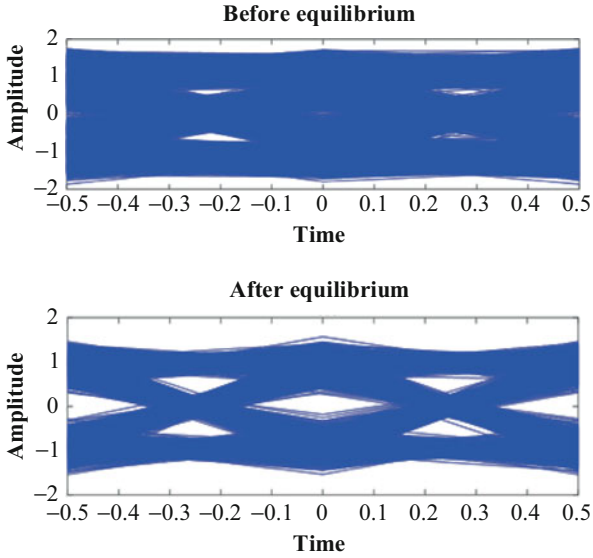


Fig. 43.8 Unequalized and equalized eye diagrams for VLC systems

performance and stability. On this basis, this study considers two typical mobile methods for indoor VLC.

Figure 43.9a shows the first movement path; the receiver PD moves from one wall corner to another wall corner in a straight line, that is, from the (0, 0, 0.65) position to the (6, 6, 0.65) position. In Fig. 43.9b, for the second movement path, the receiver PD moves along the right-angle path, that is, from the (1, 1, 0.65) position to the (5, 1, 0.65) position, and then to the (5, 5, 0.65) position.

When the receiver is moved indoors, the channel characteristics change over time, which has a significant impact on the system performance. In this method, the transmitted signal is a type of time-series data. To solve this problem, this study proposes a VLC-channel compensation method based on deep learning. The CNN-LSTM deep-learning algorithm was used in combination with the residual structure to predict the channel to compensate for the signal at the receiver and improve the performance and stability of the system.

The DCO-OFDM system was verified on a simulation platform, and the BER performance of this method under different SNRs in mobile scenes with 16QAM modulation was simulated and analyzed. Fig. 43.10a shows the trajectory of the receiver PD in space; i.e., the trajectory of the PD extending along a straight line from the starting coordinates (0, 0, 0.65) to the key coordinates (6, 6, 0.65).

Figure 43.10b shows the BER performance at a signal-to-noise ratio of $E_b/N_o =$ 10 dB. The abscissa represents the moving distance of the PD from the starting point, and the ordinate represents the bit-error rate of the system. It can be seen from the figure that the equalization method proposed in this article can effectively alleviate the distortion caused by the channel during movement. Figure 43.10c, d show the

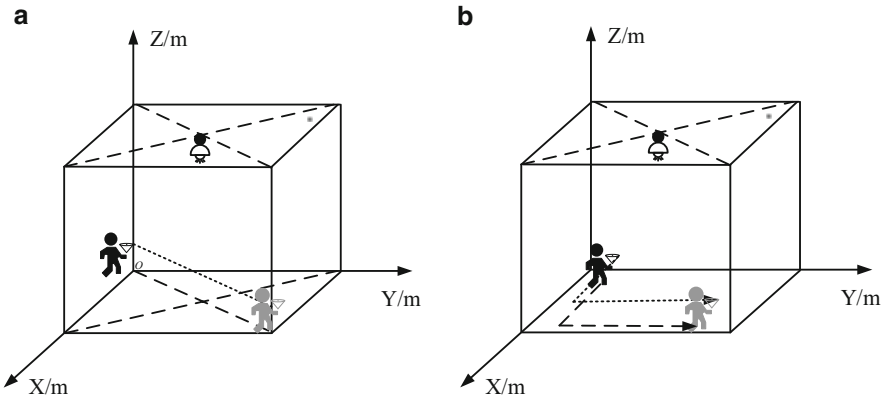


Fig. 43.9 Different movement paths in the VLC system. (**a**) Mobile path one. (**b**) Mobile path two

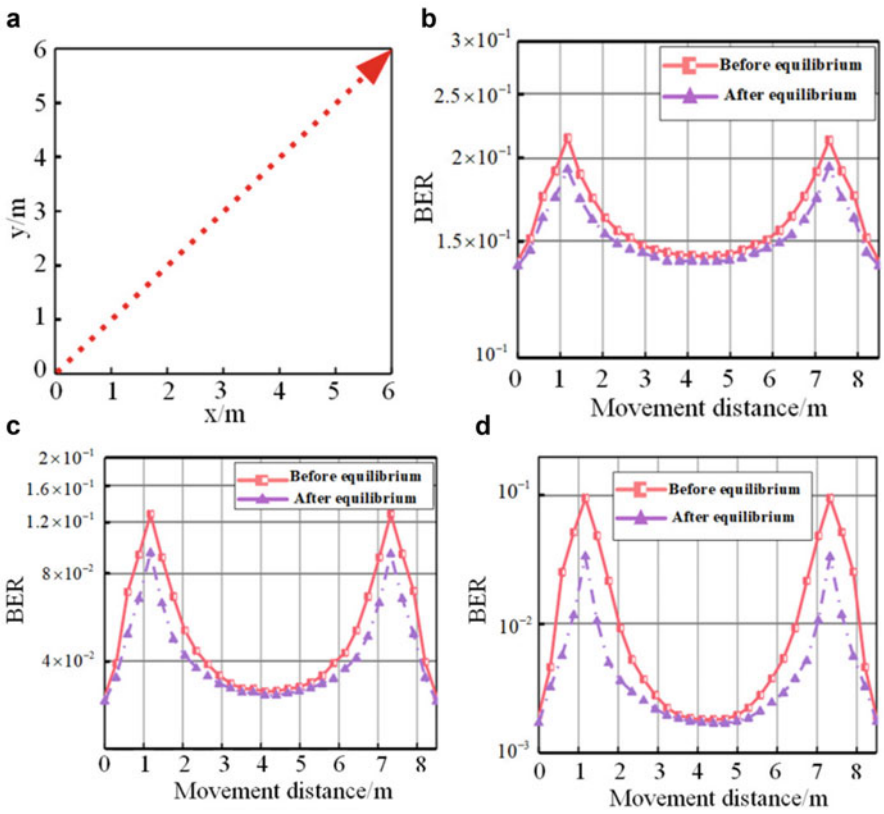


Fig. 43.10 BER performance of the DCO-OFDM system for the first mobile path. (**a**) Movement path. (**b**) $E_b/N_o = 10$ dB. (**c**) $E_b/N_o = 15$ dB. (**d**) $E_b/N_o = 20$ dB

BER performances of $E_b/N_o = 15$ dB and $E_b/N_o = 20$ dB, respectively, which also indicate that the equalization method proposed in this study has a lower bit-error rate.

Figure 43.11 shows the BER performance improvement of the DCO-OFDM system in the second mobile scenario. That is, the receiver PD moves from the starting coordinates (1, 1, 0.65) along a straight line to the coordinates (5, 1, 0.65) and then to the end point (5, 5, 0.65). The distance of the entire movement process was approximately 8 m.

Figure 43.11a shows the movement trace of the receiving PD in the spatial plane. The figure also shows that the influence of the channel on the system constantly changes during movement. The CNN-LSTM residual-structure balancing methods proposed in this article show better BER performance. Figure 43.11b shows the change in BER power as the PD moves from the start point to the end point, when $E_b/N_o = 10$ dB.

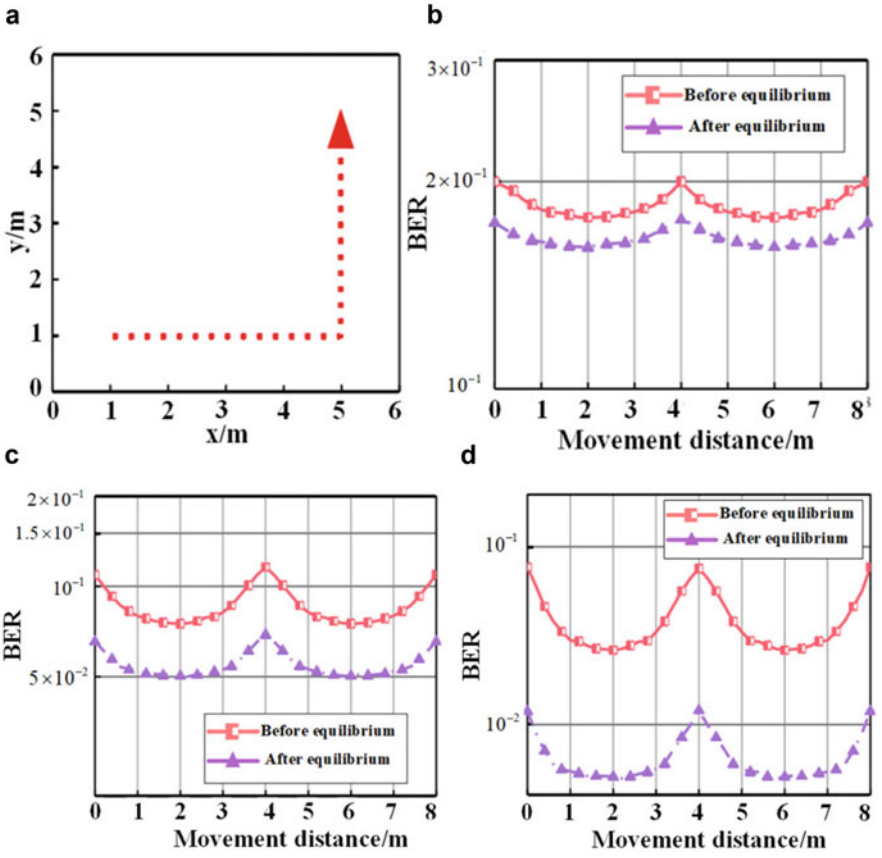


Fig. 43.11 BER performance of the DCO-OFDM system for the second path. (**a**) Movement path. (**b**) $E_b/N_o = 10$ dB. (**c**) $E_b/N_o = 15$ dB. (**d**) $E_b/N_o = 20$ dB

Figure 43.11c, d show the BER power changes at signal-to-noise ratios E_b/N_o = 15 dB and E_b/N_o = 20 dB, respectively. As shown in the figure, the CNN-LSTM residual-structure balancing method effectively alleviates the distortion caused by the channel during the entire movement process.

Comparative Simulation Analyses of Different Equalization Algorithms

To highlight the performance of the algorithm structure in this study, several existing and traditional equalization algorithms were selected for comparison. The algorithms considered in this study include the LSTM algorithm, DNN neural-network algorithm, Volterra equalization algorithm, and the proposed CNN-LSTM adaptive-equalization algorithm. The new energy of the equalization algorithm is highlighted by the change in BER under different SNRs. The BER simulation results for different algorithms are shown in Fig. 43.12, and the corresponding constellation diagrams are shown in Fig. 43.13.

Figure 43.12 shows the convergence curves of the CNN-LSTM equalizer, DNN equalizer, LSTM equalizer, and traditional Volterra blind equalizer. It can be seen from the figure that after the convergence of the CNN-LSTM equalization algorithm, the BER is much smaller than that of the LSTM, DNN, and traditional Volterra blind-equalization algorithms. This indicates that the proposed equalization algorithm has a smaller BER after convergence, higher stability, and better anti-noise performance.

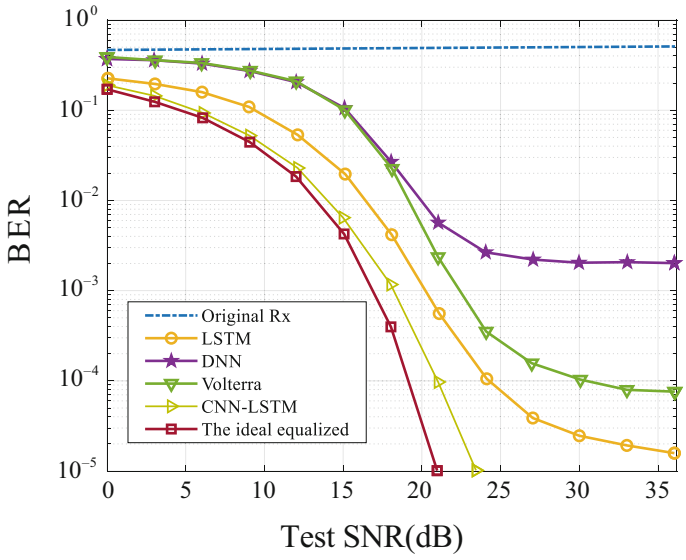


Fig. 43.12 Performance comparison of different equalization algorithms

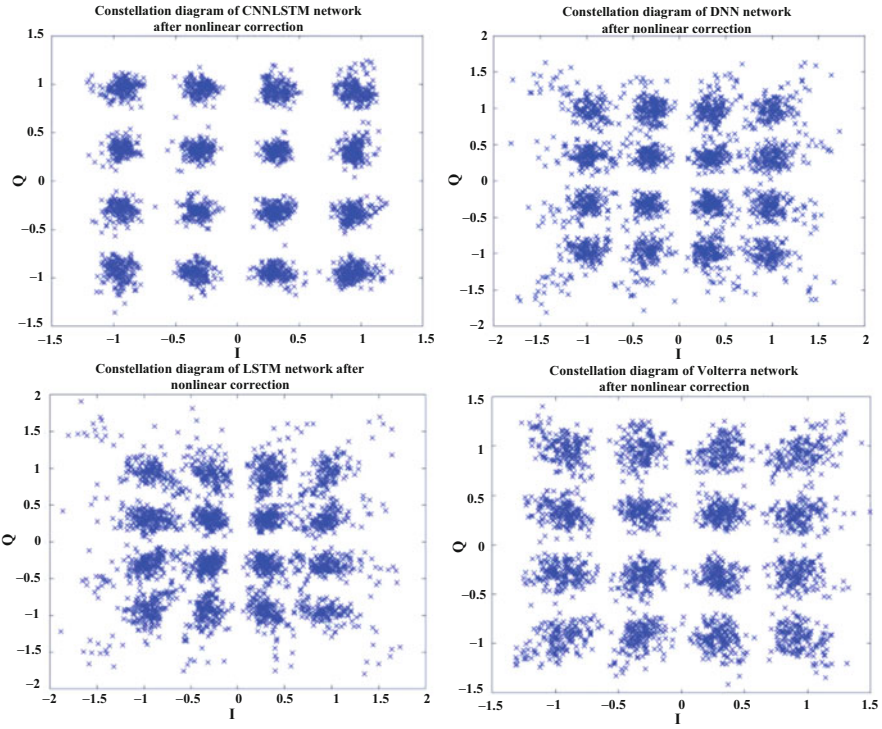


Fig. 43.13 Constellation diagrams after correction by different equalization algorithms

The correction accuracies of the DNN and LSTM algorithms converge after the signal-to-noise ratio exceeds 25 dB and the equalization performance reaches the limit. The algorithm proposed in this article is closest to the equalization curve of the ideal equalizer, does not converge, and is much better than some of the other intelligent-equalization algorithms, in terms of correction accuracy.

The effectiveness of the CNN-LSTM equalization method proposed in this article was verified using the experimental MATLAB platform. The BER performance of signals using a multipath channel, with CNN-LSTM, LSTM, DNN, and Volterra equalization in 4QAM and 16QAM DCO-OFDM systems were simulated and analyzed.

Figure 43.14 shows the BER performance of the different equalization methods in the DCO-OFDM systems. It can be observed that CNN-LSTM has the best BER performance and the same signal-to-noise ratio under both 4QAM modulation and 16QAM modulation. In contrast, the BER performances of the DNN and LSTM equalization are worse than that of the CNN-LSTM equalization, and the BER performance of the CNN-LSTM equalization is slightly better than that of LSTM equalization.

It can also be observed from the figure that as the modulation order increases, the bit-error rate increases; that is, the bit-error rate of the 16QAM modulation is higher than that of the 4QAM modulation. In addition, the equalization effect of various

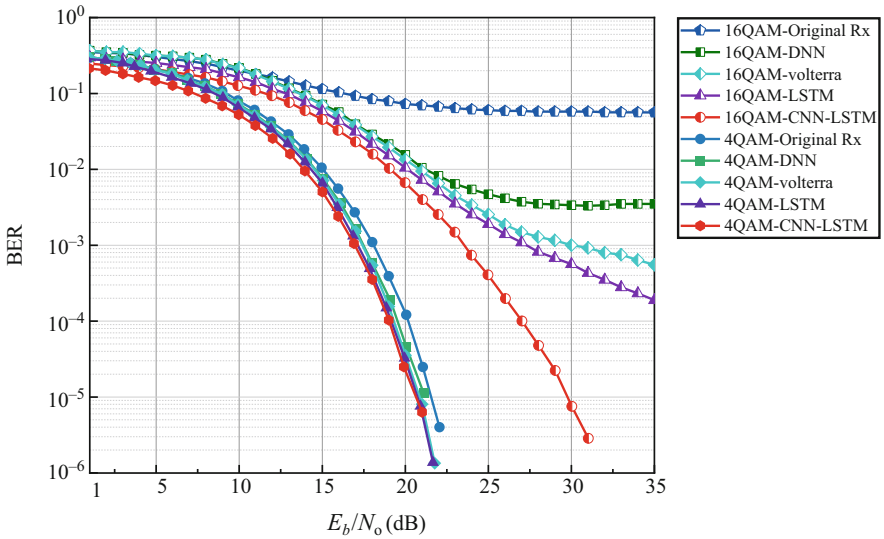


Fig. 43.14 Comparison of the BER performance of different equalization methods in a DCO-OFDM system

equalization methods is not ideal because of the strong signal distortion and high bit-error rate with a low SNR. As the signal-to-noise ratio increases, the effects of different equalization methods become apparent. In particular, the proposed CNN-LSTM method exhibits the best equilibrium effect.

Figure 45.15 shows the BER performance of the DCO-OFDM systems with 4QAM and 16QAM modulation and CNN-LSTM and LSTM equalization with different signal-to-noise ratios. As can be seen from the figure, the BER performance of the CNN-LSTM equalization method is basically the same as that of the LSTM equalization method. In addition, the BER performance of the CNN-LSTM equalization method was slightly better than that of the LSTM equalization method under different modulation orders.

Simulation Analysis of the Structural Parameters of the CNN-LSTM Algorithm

The step distance of the algorithm influences its computational speed and equalization performance. The CNN-LSTM algorithm was simulated with a fixed step size to analyze the influence of different step sizes on the bit-error rate of the algorithm. QAM modulation was adopted for the signal and the training-sequence length was 1000. The step sizes were 0.1, 0.01, 0.001, and 0.0001, and the simulation results are shown in Fig. 43.16.

As shown in Fig. 43.16, different step sizes affect the convergence speed and BER of the CNN-LSTM algorithm. From the convergence perspective, the step size

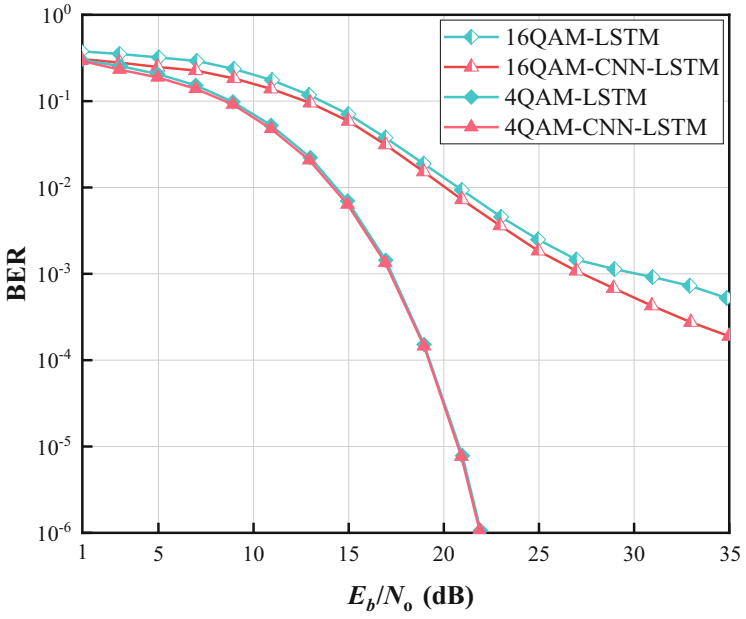


Fig. 43.15 Bit-error rate of CNN-LSTM and LSTM equalization in a DCO-OFDM system

of 0.1 converges the fastest, and a step size of 0.01 converges the next fastest. The smallest step size (0.0001) exhibits the lowest convergence.

The lowest BER was obtained when the step size was 0.0001, the next lowest was 0.001, and the highest was obtained when the step size was 0.1. From the above analysis, it can be seen that the CNN-LSTM algorithm has an inherent contradiction in the convergence speed and BER.

The distribution pattern of the LED light source and the power uniformity further increase the channel complexity, crosstalk between codes, and BER [18]. The performance of the equalizer was simulated under different light-source transparency conditions. The simulation results and constellation diagrams after balancing under the same LED transparency are shown in Figs. 43.17 and 43.18, respectively.

The transparency of the LED light source has a non-negligible influence on the performance of the equalizer: With a transparency of 0.3, the equalization residuals converge at 0.2 at an S/N ratio of 30 dB, and then stabilize. As the transparency increases, the residuals after equalization decrease and the curves converge and overlap at 0.6 and 0.8, which means that after 0.6, the transparency can be ignored and has a negligible influence on the equalizer performance.

In channel equalization, a connection exists between the channel distance and BER residuals. From Fig. 43.19, it can be seen that the BER is the lowest when the

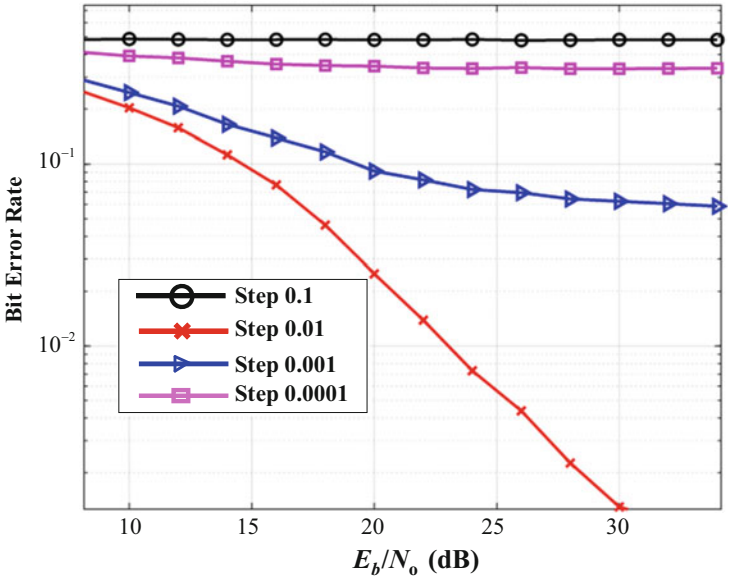


Fig. 43.16 Convergence speed as well as BER comparison of CNN-LSTM algorithms with different step sizes

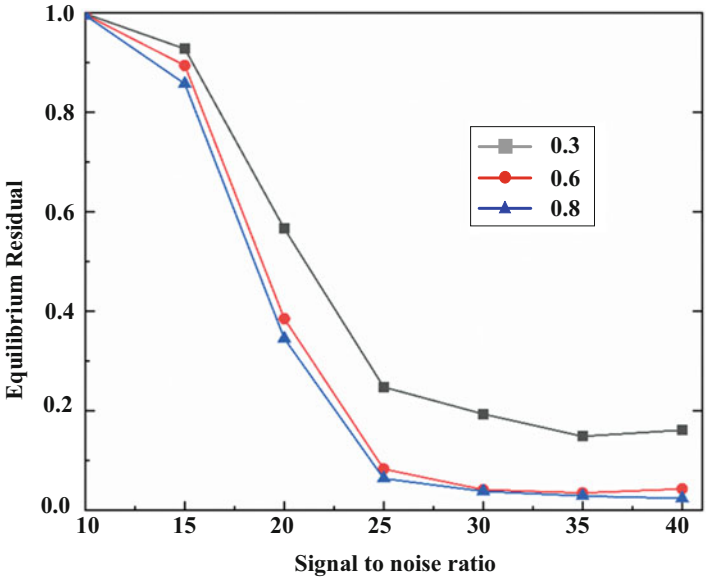


Fig. 43.17 Equalizer performance with different LED transparencies

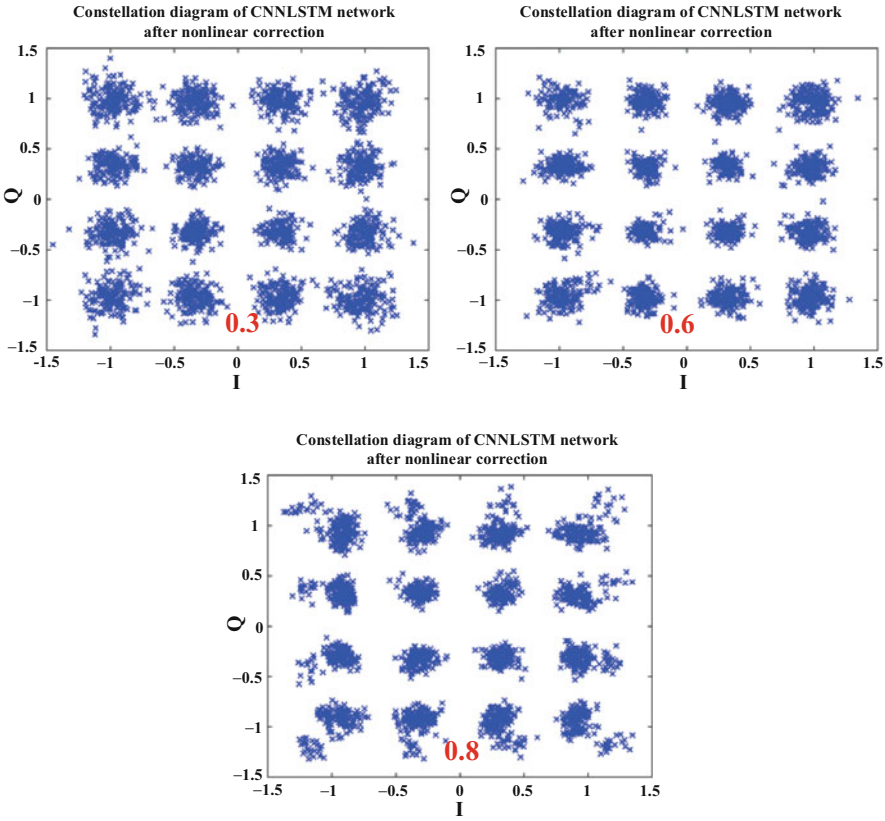


Fig. 43.18 Constellation diagrams after equalization with different LED transparencies

channel spacing is 2.0 m, whereas the BER is the highest when the channel spacing is 3.5 m. It can be clearly seen that as the channel distance increases, the BER residuals, i.e., the difference between the equalized BER and the ideal BER, increase. When the channel spacing increases from 2.0 m to 3.5 m, the BER increases; however, the equalizer always reduces the BER and meets the communication requirements after convergence (Fig. 43.20).

There is an inevitable relationship between the BER residuals and the signal-current threshold. From Figs. 43.21 and 43.22, we can see that as the signal-current threshold increases, the BER residuals decrease; however, the balanced residuals converge at an SNR of 20 dB. Then, the current threshold increases and the BER residuals tend to stabilize; hence, the SNR is optimal at 20 dB. It also shows that increasing the signal current improves the performance of the system and the quality of the signal (Fig. 43.23).

To check the robustness of the CNN-LSTM algorithm, random interference was artificially set into the channel with a certain threshold, according to the method

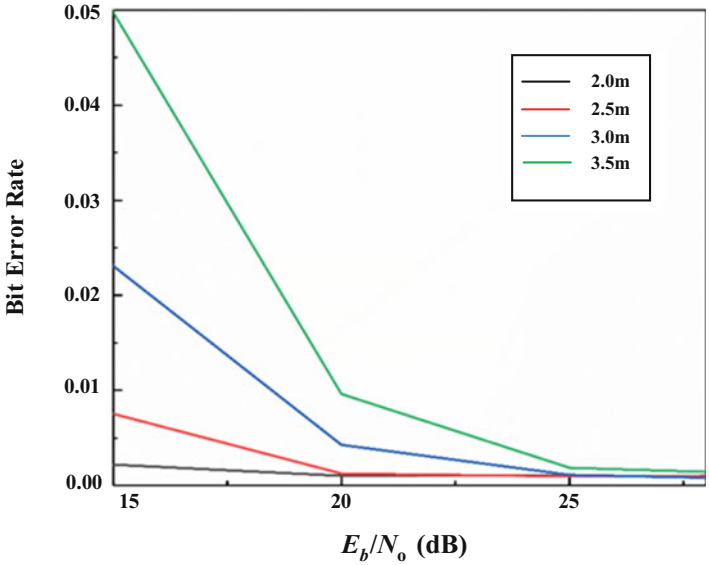


Fig. 43.19 BER residuals after equalization for different channel spacings

described in section "Robustness Analysis of the Algorithm." The simulation results showed that CNN-LSTM is more robust than the LSTM, DNN, and Volterra equalization algorithms.

Conclusion

In this paper, we proposed a convolutional neural network (CNN) fused with an LSTM time-series memory structure for a visible-light communication equalization algorithm. The CNN learns the channel features and converts them into a vector of tapped coefficients, which is described as a time-series feature and stored in the LSTM time-series memory structure. The structure can actively forget and wake up the CNN structure to re-learn the channel features to adjust the equilibrium when the channel is changed.

In this study, the channel-compensation performance of the equalizer was analyzed by numerical simulation. The simulation results indicated the following:

1. In the 35-dB signal-to-noise ratio, the receiver side of the BER reduced below 10^{-3}, to meet the communication requirements.
2. The mapped amplitude after equalization fit well with the ideal accepted amplitude, indicating that the accuracy of the equalization algorithm had been further improved. The widening of the eye diagram also indicated that the communication quality had been further improved.

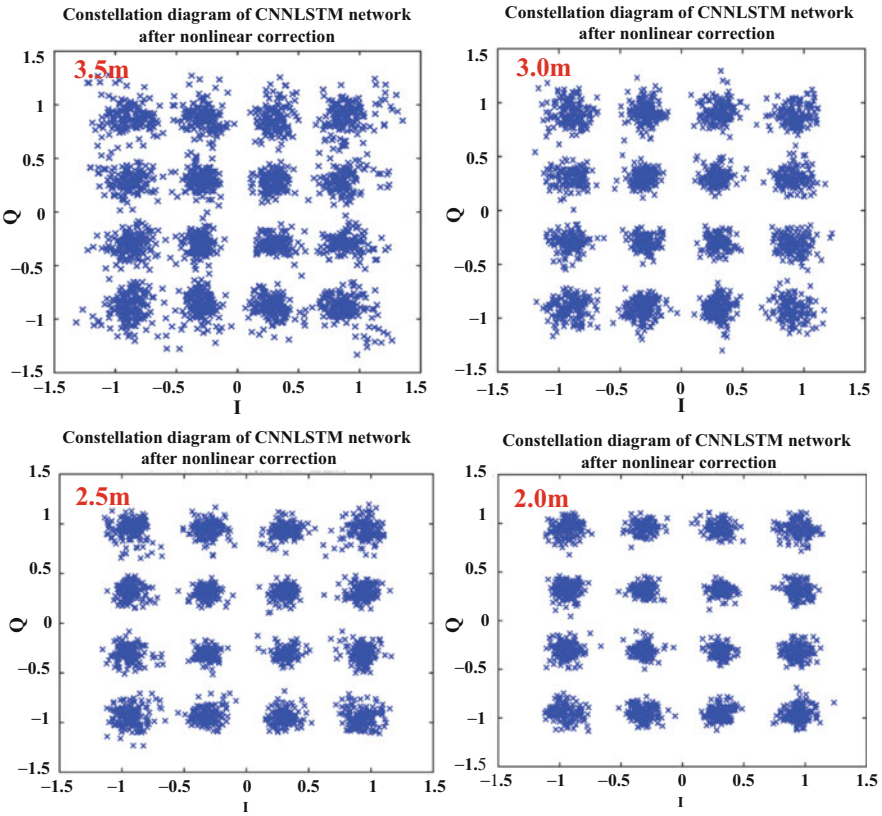


Fig. 43.20 Constellations of received signals at different channel distances

3. In the comparative analysis with existing algorithms, the proposed algorithm had the highest equalization accuracy and actively learned the parameters for channel changes and adaptive equalization.
4. The simulation results under different VLC parameters showed that the method was beneficial for improving the overall characteristics of the channel. A better BER performance was obtained at an affordable complexity cost, with an improvement of at least 2.8-dB SNR over the correction accuracy of the conventional method at a BER of 10^{-3}.

To solve the problem of channel characteristics changing with time with a mobile receiver in an indoor VLC system, this study considered two typical mobile-path models and analyzed the influence of interference factors, such as channel noise and multipath impact, on the BER performance of the system. Considering this scenario, a channel-compensation method based on the CNN-LSTM residual algorithm was proposed to reduce the influence of the channel on the system performance when the receiver moved. The simulation results showed that the proposed method could

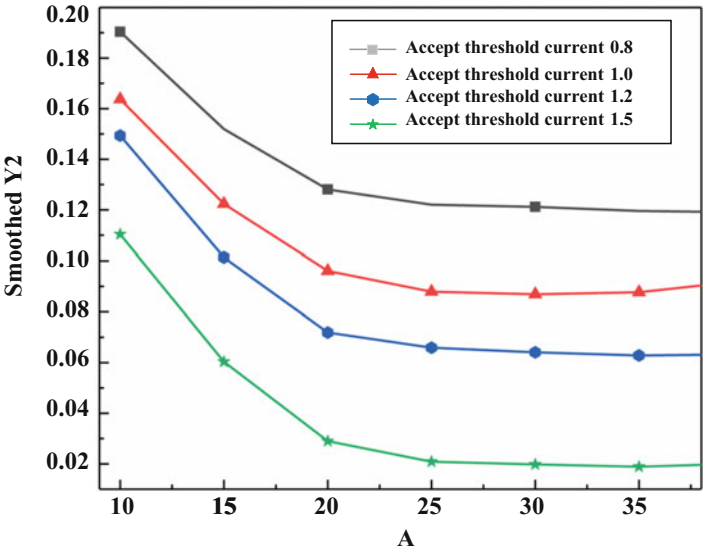


Fig. 43.21 BER residuals after equalization of signal currents with different thresholds

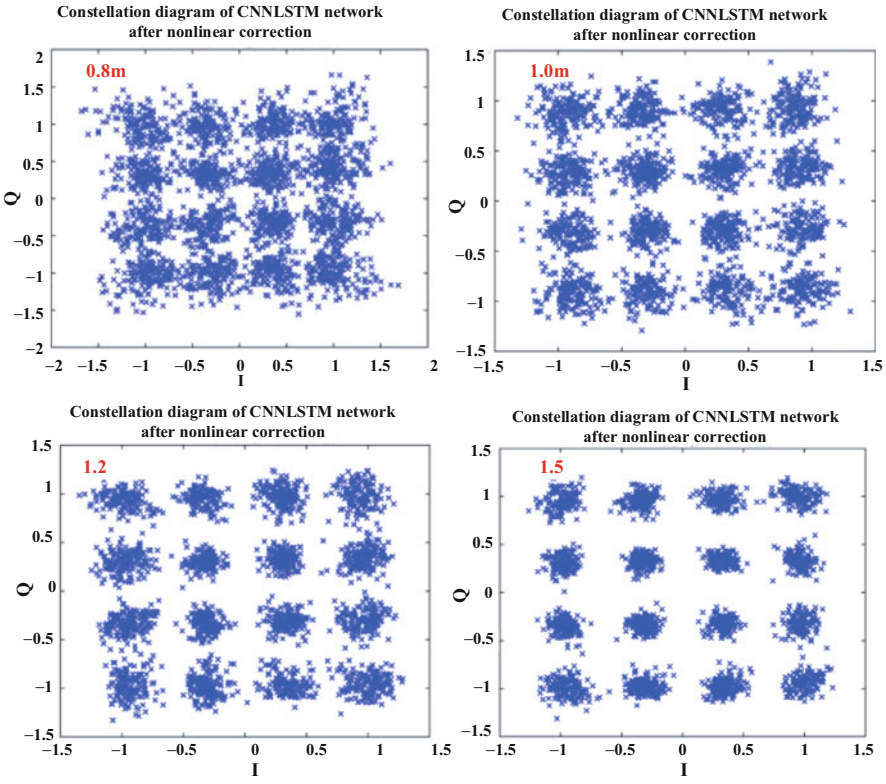


Fig. 43.22 Constellation diagrams at different signal thresholds

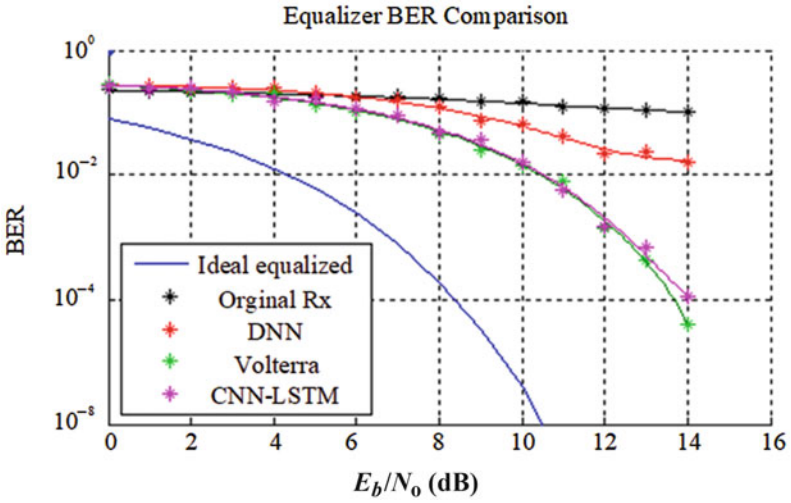


Fig. 43.23 Performance comparison of different equalization algorithms with random interference

effectively contain the signal distortion caused by the channel and improve the reliability and stability of the system.

References

1. Chen jinguang. Shot noise and Thermal noise in PN Junction Diode [J]. Acta Phys. Sin, 1965, 21(2): 383–389.
2. Fan K, Komine T, Tanaka Y, et al. The Effect of Reflection on Indoor Visible Light Communication System Utilizing White LEDs [C]. The 5th International Symposium on Wireless Personal Multimedia Communication, Honolulu, HI, USA, 2002(2): 611~615
3. Lu Xingyu, Lu Chao, Yu Weixiang, et al. Memory-controlled deep LSTM neural network post-equalizer used in high-speed PAM VLC system [J]. Optics express, 2019, 27(5): 7822–7833.
4. Lin Bangjiang, Lai Qiwei, Ghassemlooy Z., et al. A machine learning based signal demodulator in NOMA-VLC [J]. Journal of Lightwave Technology, 2021, 39(10): 3081–3087.
5. Zhao Yiheng, Zou Peng, He Zhixue, et al. Low spatial complexity adaptive artificial neural network post-equalization algorithms in MIMO visible light communication systems [J]. Optics Express, 2021, 29(20): 32728–32738.
6. Lu Xingyu, Liu Yuanyuan, Chen Junjie, et al. Patterns quantization with noise using Gaussian features and Bayesian learning in VLC systems [J]. Journal of Lightwave Technology, 2022.
7. Chen Hui, Jia Junlian, Niu Wenqing, et al. Hybrid frequency domain aided temporal convolutional neural network with low network complexity utilized in UVLC system [J]. Optics Express, 2021, 29(3): 3296–3308.
8. Ohno S, Teo K a D. Tradeoff between BER and Bandwidth Efficiency in MIMO Systems with ZF Equalization [C] //2006 International Symposium on Intelligent Signal Processing and Communications. 2006: 518–521.
9. Beex A, Ikuma T. Bi-scale LMS equalization for improved performance [C] //2008 IEEE 9th Workshop on Signal Processing Advances in Wireless Communications. 2008: 511–515.

10. Wang L, He W, Zhou K, et al. Adaptive channel equalization based on RLS algorithm [C] // 2011 International Conference on System science, Engineering design and Manufacturing informatization. 2011: 105–108.
11. Ju M, Li L. A simple simi-blind equalization algorithm for precoding single carrier system with frequency domain equalization [C] //2010 IEEE International Conference on Information Theory and Information Security. 2010: 1019–1021.
12. Wolf M, Grobe L, Rieche M R, et al. Block transmission with linear frequency domain equalization for dispersive optical channels with direct detection [C] //2010 12th International Conference on Transparent Optical Networks. 2010: 1–8.
13. Xie xiuxiu, Jiao yuan, Cao lingling, et al. Linear Equalization and decision feedback Equalization in Indoor visible Light Communication [J]. Informatization Research, 2011, 37(002): 29–32.
14. Komine T, Lee J H, Haruyama S, et al. Adaptive equalization system for visible light wireless communication utilizing multiple white LED lighting equipment [J]. IEEE Transactions on Wireless Communications, 2009, 8(6): 2892–2900.
15. Bandara K, Chung Y H. Reduced training sequence using RLS adaptive algorithm with decision feedback equalizer in indoor Visible Light Wireless Communicationchannel [C] // 2012 International Conference on ICT Convergence (ICTC). 2012: 149–154.
16. Chen C, Zhong W D, Wu D. Indoor OFDM visible light communications employing adaptive digital pre-frequency domain equalization [C] //2016 Conference on Lasers and Electro-Optics (CLEO). 2016: 1–2.
17. Jin B, Zhang M, Zhang Y, et al. Visible light communications using blind equalization [C] // 2011 Asia Communications and Photonics Conference and Exhibition (ACP). 2011: 1–6.
18. Cao P, Chen J, You X. An initialization scheme for blind equalization in VLC systems [C] // 2017 16th International Conference on Optical Communications and Networks (ICOCN). 2017: 1–3.
19. Shailaja K, Seetharamulu B, Jabbar M A. Machine learning in healthcare: A review [C] //2018 Second international conference on electronics, communication and aerospace technology (ICECA). 2018: 910–914.
20. Ahmed S, Alshater M M, El Ammari A, et al. Artificial intelligence and machine learning in finance: A bibliometric review [J]. Research in International Business and Finance, 2022, 61: 101646.
21. Maschler B, Weyrich M. Deep transfer learning for industrial automation: a review and discussion of new techniques for data-driven machine learning [J]. IEEE Industrial Electronics Magazine, 2021, 15(2): 65–75.
22. Sun H, Zhang Y, Wang F, et al. SVM aided signal detection in generalized spatial modulation VLC system [J]. IEEE Access, 2021, 9: 80360–80372.
23. Gao Z, Wang Y, Liu X, et al. FFDNet-Based Channel Estimation for Massive MIMO Visible Light Communication Systems [J]. IEEE Wireless Communications Letters, 2020, 9(3): 340–343.
24. Turan B, Coleri S. Machine learning based channel modeling for vehicular visible light communication [J]. IEEE Transactions on Vehicular Technology, 2021, 70(10): 9659–9672.
25. Knudde N, Raes W, Bruycker J D, et al. Data-Efficient Gaussian Process Regression for Accurate Visible Light Positioning [J]. IEEE Communications Letters, 2020, 24(8): 1705–1709.
26. Xiao ying, Li zhenxing, Liu guozhi, et al. Research on variable step size Neural Network Blind Equalization algorithm in underwater Acoustic Communication [J]. Technical A coustics, 2005, 24(3): 129–131+136.
27. Haigh P A, Ghassemlooy Z, Rajbhandari S, et al. Visible Light Communications: 170 Mb/sUsing an Artificial Neural Network Equalizer in a Low Bandwidth White Light Configuration [J]. Journal of Lightwave Technology, 2014, 32(9): 1807–1813.

28. Fuertes G, Navarrete R, Alfaro M, et al. Adaptive equalization using artificial neural networks for a visible light communication system [C] //2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA). 2018: 1–6.
29. Zhang J, Gao M, Chen W, et al. Non-Data-Aided k-Nearest Neighbors Technique for Optical Fiber Nonlinearity Mitigation [J]. Journal of Lightwave Technology, 2018, 36(17): 3564–3572.
30. Wang D, Zhang M, Cai Z, et al. Combatting nonlinear phase noise in coherent optical systems with an optimized decision processor based on machine learning [J]. Optics Communications, 2016, 369: 199–208.
31. Ma J, He J, Shi J, et al. Nonlinear Compensation Based on K-Means Clustering Algorithm for Nyquist PAM-4 VLC System [J]. IEEE Photonics Technology Letters, 2019, 31(12): 935–938.
32. Lu X, Zhao M, Qiao L, et al. Non-linear compensation of multi-CAP VLC system employing pre-distortion base on clustering of machine learning [C] //Optical Fiber Communication Conference. 2018: M2K. 1.
33. Weixiang Y, Xingyu L, Nan C. Signal decision employing density-based spatial clustering of machine learning in PAM-4 VLC system [C] //ProcSPIE. 2018: 295–299.
34. Miao P, Yin W, Peng H, et al. Deep Learning based Nonlinear Equalization for DCO-OFDM Systems [C] //2021 IEEE International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT). 2021: 699–703.
35. Liu H, Yang X, Chen P, et al. Deep Learning Based Nonlinear Signal Detection in Millimeter-Wave Communications [J]. IEEE Access, 2020, 8: 158883–158892.
36. Li G, Hu F, Zhao Y, et al. Enhanced Performance of a Phosphorescent White LED CAP 64QAM VLC system utilizing Deep Neural Network (DNN) Post Equalization [C] //2019 IEEE/CIC International Conference on Communications in China (ICCC). 2019: 173–176.
37. Costa W S, Samatelo J L A, Rocha H R O, et al. Direct Equalization with Convolutional Neural Networks in OFDM based VLC Systems [C] //2019 IEEE Latin-American Conference on Communications (LATINCOM). 2019: 1–6.
38. Játiva P P, Becerra R, Azurdia-Meza C A, et al. Extreme Learning Machine Based Channel Estimator and Equalizer for Underground Mining VLC Systems [C] //2021 IEEE Latin-American Conference on Communications (LATINCOM). 2021: 1–6.
39. Song ziyu. Research and implementation of visible Light Communication system [D]. Fujian: Huaqiao University, 2017.
40. Ke xizheng, Qin huanhuan, Yang shangjun. et al. Night background light noise model of visible light communication system in vehicle network [J]. Chinese Journal of Radio Science, 2021, 36(06):986–990.
41. Y. Li, M. Chen, Y. Yang, M. -T. Zhou and C. Wang, "Convolutional recurrent neural network-based channel equalization: An experimental study," 2017 23rd Asia-Pacific Conference on Communications (APCC), Perth, WA, Australia, 2017, pp. 1–6, doi: https://doi.org/10.23919/APCC.2017.8304090.
42. G. Kechriotis, E. Zervas and E. S. Manolakos, "Using recurrent neural networks for adaptive communication channel equalization," in IEEE Transactions on Neural Networks, vol. 5, no. 2, pp. 267–278, March 1994, doi: https://doi.org/10.1109/72.279190.
43. LeCun Y. LeNet-5, convolutional neural networks [J]. URL: http://yann.lecun.com/exdb/lenet, 2015, 20(5): 14.
44. Li X, Yi X, Liu Z, et al. Application of novel hybrid deep leaning model for cleaner production in a paper industrial wastewater treatment system [J]. Journal of Cleaner Production, 2021(2): 126343.
45. He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification [C] //Proceedings of the IEEE international conference on computer vision. 2015: 1026–1034.

46. Lan P, Xia K, Pan Y, et al. An improved equilibrium optimizer algorithm and its application in LSTM neural network [J]. Symmetry, 2021, 13(9): 1706.
47. Kingma D P, Ba J. Adam: A method for stochastic optimization [J]. arXiv preprint arXiv:1412.6980, 2014.
48. Yiheng Zhao, Peng Zou, Zhixue He, Ziwei Li, and Nan Chi, "Low spatial complexity adaptive artificial neural network post-equalization algorithms in MIMO visible light communication systems," Opt. Express 29, 32728–32738 (2021)

Ultraviolet Non-Line-of-Sight Communication

44

Contents

Introduction 1620
Research Progress at Home and Abroad 1620
Foreign Research Progress in Ultraviolet Optical Communication 1620
Domestic Research Progress in Ultraviolet Optical Communication 1622
Channel Model 1623
Single-Scattering Model 1624
Multiple-Scattering Model 1625
UV NLOS Communication Link Model 1625
Pulse-Broadening Effect of NLOS Ultraviolet Communication 1628
Influence of Atmospheric Turbulence on Ultraviolet NLOS Communication 1630
Key Technologies of UV Self-Organizing Networks 1632
Node-Localization Algorithm of an Ultraviolet Optical Communication Network 1633
Ultraviolet Self-Organizing-Network Multiuser Detection Technology 1634
MAC-Layer Protocol 1637
UV Ad hoc: Network Routing Protocol 1637
Analysis of Energy Saving in UV Wireless Sensor Networks 1641
Summary 1642
References 1644

Abstract

Ultraviolet technology has been widely used in military and civil fields, such as space detection, ultraviolet guidance, ultraviolet interference, corona discharge detection, and military communication. In this paper, the development of UV communication at home and abroad is introduced from the single and multiple scattering of UV light, the different links in UV communication are analyzed, the pulse broadening effect of non-direct view link UV communication and the influence of atmospheric turbulence are discussed, the UV self-organizing network is summarized, and the development of UV communication is also looked forward.

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_44

Keywords

Ultraviolet optical communication · Channel modeling · Pulse broadening · Ad hoc networks

Introduction

Ultraviolet communication is gradually becoming a research hotspot because of its inherent advantages such as low solar background noise, non-direct vision, high confidentiality, and good concealment. The strong scattering characteristics of UV band in the atmosphere make it an ideal choice for Non-Line-of-Sight (NLOS) communication. And because the transmission process is affected by Rayleigh scattering and Mie scattering, the UV light intensity decays rapidly as the transmission distance increases, so the UV light signal can only be received within a certain range, making its communication difficult to be eavesdropped. UV light communication is usually short-distance communication under complex weather conditions, although the strong absorption of the atmosphere, but due to the haze weather, dust, battlefield, and other environmental aerosol density increases, will strengthen the scattering ability of UV light, but to improve the stability of the UV light system communication link.

Research Progress at Home and Abroad

Owing to its strong scattering properties in the atmosphere, the ultraviolet band is ideal for non-line-of-sight (NLOS) communication [1]. Because the transmission process is affected by Rayleigh scattering and Mie scattering, with the increase in transmission distance, the ultraviolet light intensity is rapidly attenuated; therefore, ultraviolet light signals can be received only within a certain range, thus avoiding eavesdropping. Ultraviolet communication is usually short-distance communication under complex weather conditions. Because of the increase in aerosol density owing to hazy weather, dust, battlefields, and other environments, the scattering ability of ultraviolet light is enhanced, and the stability of communication links of ultraviolet light systems is enhanced [1].

The ozone layer in the stratosphere has strong absorption characteristics for the 200–260 nm ultraviolet light band. Hence, the ultraviolet light in this band has very low radiation when it reaches the surface of the earth, which is called the solar-blind zone. The solar-blind zone creates a superior communication condition for communication systems working in this band [1].

Foreign Research Progress in Ultraviolet Optical Communication

In 1939, the US government discussed the Navy's ultraviolet communication system, including light sources, photodetectors, and ultraviolet filters, but did not conduct experiments [2].

In 1968, Sunstein et al. [3] built an optical communication link and provided experimental ultraviolet results for the system under different parameters. In 1976, under the new Standard Atmosphere model, Reilly proposed and developed a theoretical model for describing single scattering, based on ellipsoidal coordinate systems, and evaluated various combinations of absorption and scattering coefficients, anisotropic phase functions, and transmitter/receiver distances [4].

In 1985, Geller et al. [5] used a mercury lamp as an ultraviolet light source and a photomultiplier tube and a filter as a receiver to verify the feasibility of ultraviolet communication at a communication distance of 1 km. In 1991, Luettgen et al. [6] proposed a single-scattering model describing uniform scattering and dielectric absorption, and analyzed the pulse response time and path loss of short-range indirect communication systems.

In 2000, the General Telephone & Electronics (GTE) company, with funding from the US military, developed an indirect ultraviolet optical communication system with a 4.8-kbit/s transmission rate at a communication distance of less than 2 km, and a bit-error rate of 10^{-6}, which met the requirements for video transmission [7].

In 2005, Shaw et al. [8] conducted field experiments in various environments, such as open areas and under canopies. Measurements showed that solar-blind UV light-emitting diodes (LEDs) could be applied to short-range communication links.

In 2008, Xu et al. [9] used on–off keying (OOK) modulation in a communication system. The bit-error rate for different transmission elevation angles was obtained by changing the transmission and receiving angles. When the emission elevation angle was 30°, the bit-error rate increased from 10^{-6} to 10^{-1}. In 2011, Xu et al. [10] compared the path-loss and bit-error rate performance of links in different environments. In 2013, they proposed a configurable framework that combined the traditional line-of-sight turbulence theory with a Monte Carlo simulation framework for random photon scattering [11].

In 2015, Mohamed et al. [12] applied delay-compensation and spatial-diversity techniques to suppress detector noise in indirect UV optical communications. Experiments showed that under the same signaling conditions, the information-transmission rate of diversity technology was 50% higher than that of a single receiver.

In 2016, Ardakani et al. [13] analyzed the relationship between the signal-to-noise ratio and bit-error rate in three different modes: NLOS ultraviolet multiple-input multiple-output (MIMO), multiple-input single-output (MISO), and single-input multiple-output (SIMO).

In 2018, Refaai et al. [14] used decoding-and-forwarding relay technology to complete NLOS UV communication according to the channel-state information (CSI). The interruption probability of a multi-relay UV system was discussed, and the system performance at different elevation angles, atmospheric scattering parameters, and receiver fields of view were simulated and analyzed.

In 2019, Arya et al. [15] proposed a multi-hop ultraviolet indirect vision communication system based on amplification and forwarding and adopted subcarrier-intensity modulation to suppress the influence of turbulence. Simulation results showed that the system could effectively mitigate turbulence-induced fading without

sacrificing important performance criteria, such as the interrupt probability and average bit-error rate.

In 2020, Hasan et al. [16] derived a system model under MISO, SIMO, and MIMO modes by studying the spatial-diversity reception technology in NLOS ultraviolet optical communication. They analyzed the bit-error rate performance of the communication link in a single-input single-output (SISO) system model.

In 2021, Bhatia et al. [17] proposed an expression for the interrupt probability of a nonlinear ultraviolet communication system and proposed an equalizer based on the minimum bit-error rate criterion of the reproducing kernel Hilbert space (RKHS) to reduce the distortion generated by LED nonlinearity.

In 2022, Arslan et al. [18] proposed a steering optimization method based on a finite difference stochastic approximation (FDSA). It could optimize the pointing direction of both the transmitter (T_x) and receiver (R_x) without knowing the position or relative orientation of the two nodes.

Domestic Research Progress in Ultraviolet Optical Communication

Major domestic research institutions include the Beijing Institute of Technology, Xi'an University of Technology, Chinese Academy of Sciences, and other institutions. Most studies are still in the prototyping and feasibility testing stages, and some research on channel models is being explored.

In 1999, Ni et al. [19] achieved voice communication at a distance of 500 m using a low-pressure-gas ultraviolet mercury lamp with a 253-nm wavelength as the light source. In 2003, the team analyzed the influence of factors, such as elevation difference, visibility, weather, and transmission distance, on the atmospheric ultraviolet transmittance at the transmitter and receiver ends [20].

In 2005, Lan et al. [21] proposed to use frequency modulation for language coding in ultraviolet-light information transmission. A data-transmission rate of 2 kbit/s laid a good foundation for the solar-blind ultraviolet communication system.

In 2007, Xiao [22] used an ultraviolet LED to replace a low-pressure mercury lamp and adopted pulse-position modulation (PPM) technology to realize an experimental prototype with a 10-m communication distance.

In 2010, Ke [23] proposed to use the multi-hop feature of self-organizing networks to overcome the limited ultraviolet transmission distance, and added a reward-and-punishment mechanism to improve the original binary exponential-retreat algorithm. The competitiveness of the nodes is matched with their competitive windows, thereby achieving fair access to channels.

In 2011, considering the indirect UV optical communication characteristics, the coverage azimuth was characterized and the NLOS directional transmission node was improved. Moreover, a directional medium-access control (MAC) protocol for indirect UV optical communication based on angle perception was proposed and simulated [24]. The protocol not only shortened the average communication delay but also increased the network throughput. In 2012, a Monte Carlo (MC) multi-

scattering ultraviolet optical-channel model was continuously studied. Based on the single-scattered channel model, the node coverage of short-range communication was analyzed, and its mathematical expression was derived [25].

In 2013, Wu et al. [26] proposed a MAC-layer protocol suitable for ultraviolet optical communications. In the same year, Han et al. [27] adopted a low-density parity-check code algorithm to improve the system performance, and improved the communication distance without coding by 78% under the rated transmit power and a bit-error rate of 10^{-5}.

In 2015, Yang et al. [28] proposed a UV directional neighbor-discovery algorithm based on traditional time-division multiple access (TDMA), which can be applied to large-scale UV ad hoc networks with high-speed movement. Compared with the traditional algorithm, the probability of neighbor-node discovery was increased by 45%, and node collisions were reduced; however, the algorithm convergence speed was sacrificed.

In 2016, Song et al. [29] used the MC method to simulate and analyze the interference between links in space-division multiplexing and noted the pulse-broadening problem in non-coplanar conditions. In 2018, a UV mobile ad hoc communication node device was designed, and transmission and reception were integrated [30].

In 2019, Wang et al. [31] proposed the use of multi-decimal frequency-shift keying for modulation in view of the low modulation rate, and developed an ultraviolet optical communication system for experiments. Using quaternary frequency-shift keying (QFSK), the maximum transmission rate reached 30 kbit/s, which was double that of the binary FSK method.

In 2020, Zhang et al. [32] built a field-programmable gate array (FPGA)-based ultraviolet optical communication system, and analyzed its time delay. The results showed that using an FPGA to build the communication system effectively reduced the communication delay.

In 2021, Cao et al. [33] proposed a single-scattering loss model for an NLOS UV communication system based on an ultraviolet LED. Compared with the actual measurement results, the root-mean-square error of the proposed loss model was only about 1 dB.

In 2022, Cao et al. [34] established a reflection-assisted NLOS ultraviolet communication path-loss analysis model combined with a single collision-induced event (CIE), which worked well in both coplanar and non-coplanar geometries of reflection-assisted NLOS UV communication.

Channel Model

In short-range (usually less than 200 m) ultraviolet communications, the general single-scattering model is a good representation of the channel model. For medium- and long-range ultraviolet communications, the Monte Carlo method multiple-scattering model is more consistent with the experimental results [35].

Single-Scattering Model

As shown in Fig. 44.1, the emitter emits a beam with divergence and elevation angles. A receiver with an angle of view and elevation detects the scattered light from the atmosphere. The linear distance between the receiving and transmitting ends is r, and the common volume of the angle of view and beam angle is V.

In Fig. 44.1, F_1 and F_2 are the focus positions in the ellipsoidal coordinates. $2\Phi_t$ and $2\Phi_r$ represent the divergence angle at the beginning and the field angle at the end, respectively. The emitter elevation angle is relative to the horizontal plane; θ_r represents the angle of view of the receiver, and β_s is the scattering angle.

Feng [36] simulated the communication at different transmit and receive angles under different transmission distances. When the transmission distance increased from 1 m to 1 km and the bit-error rate increased from 10^{-10} to 10^{-1}, the simulation results are shown in Fig. 44.2. It can be observed that with an increase in distance, the overall transmission rate shows a downward trend.

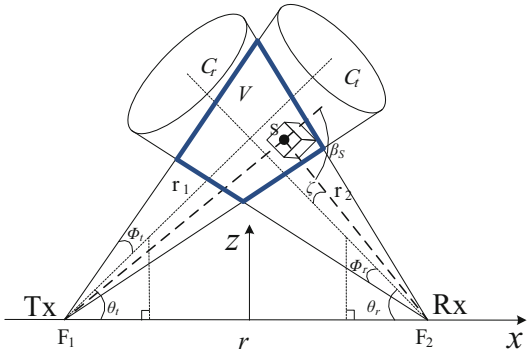


Fig. 44.1 Single-scattering model [1]

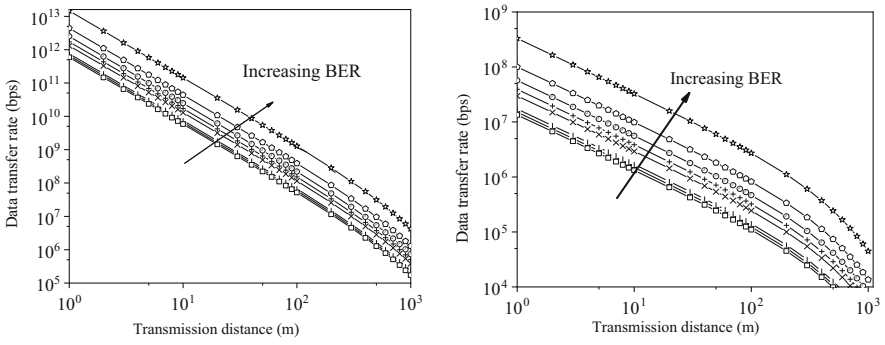


Fig. 44.2 Data-transfer rate comparison between LOS and NLOS systems [36]

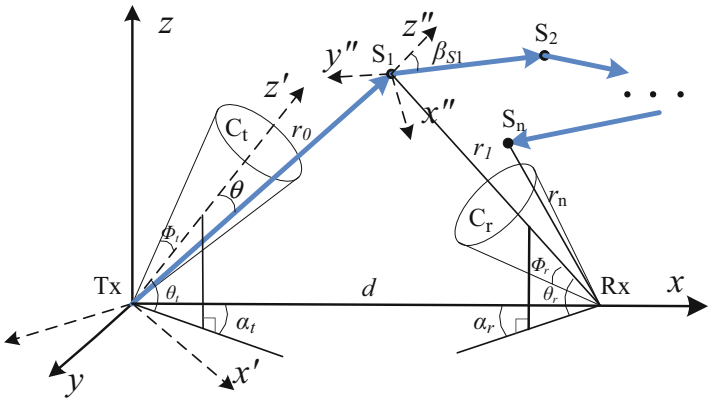


Fig. 44.3 Multiple-scattering propagation model [39]

Multiple-Scattering Model

As shown in Fig. 44.3, for the multiple-scattering model, to simplify subsequent coordinate calculations, T_x was placed at the origin (0, 0, 0) and R_x at the position of point (d, 0, 0) [38]. $2\Phi_t$ denotes the divergence angle of the beam at the emitting end, α_t denotes the offset angle of the emitting end, and α_r denotes the offset angle of the receiving end. $2\Phi_r$ denotes the angle of view of the detector at the receiving end; θ_t and θ_r denote the elevations of the starting and receiving ends, respectively; r_0 and r_1 denote the distance from T_x to the scattering point S_1 and from the scattering point S_1 to R_x, respectively.

Song [39] simulated the path loss of a UV NLOS communication system using the ergodic micro-element and pointing-probability methods. The simulation results are shown in Fig. 44.4. It can be seen from the graph that the path loss tends to increase and the common scatterers decrease as the offset angle α_r of the receiver increases.

UV NLOS Communication Link Model

Ultraviolet-light NLOS communication means that in the transmission process, the receiving and sending ends do not need to be aligned. It relies on ultraviolet-light scattering by atmospheric molecules to carry out barrier communication. The scattering occurs at the overlap of the transmitting and receiving fields of view, which is called the effective scatterer. According to the difference between the transmit and receive elevations, Ke [1] divided the scattering into three communication modes, as shown in Fig. 44.5a–c.

Different NLOS links can be used in different communication situations because of the different transceiver-elevation angles and communication distances. See Ref. [37] for a specific comparison. In Fig. 44.6, both ends are at the same level.

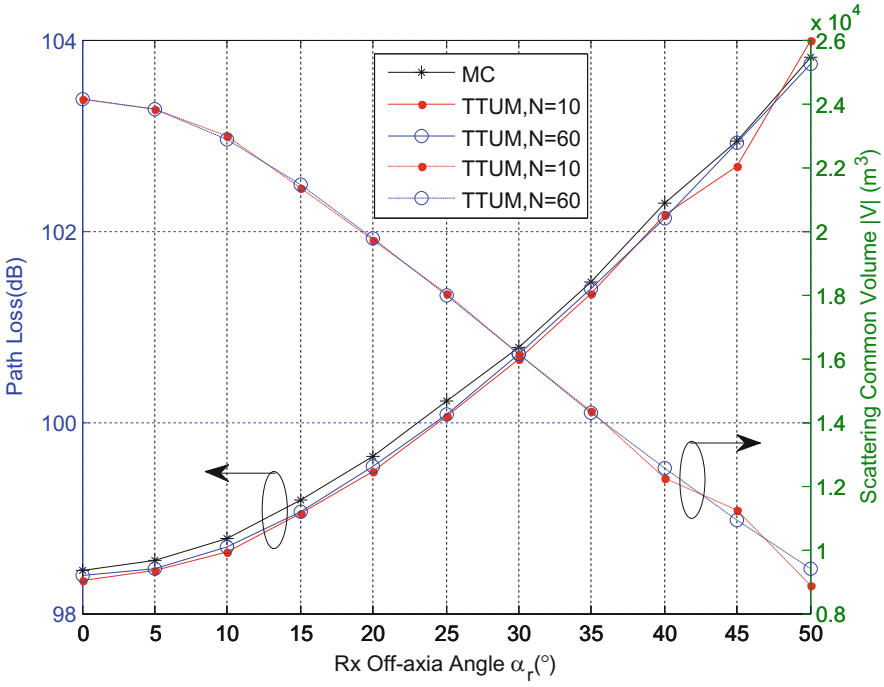


Fig. 44.4 Path loss at different receiver-offset angles [39]

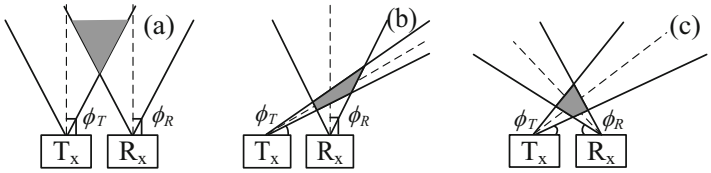


Fig. 44.5 Typical NLOS communication links [1]

The transmitting end emits a beam with divergence angle Φ_1 and elevation angle θ_1. The receiving end detects the scattered light from the atmosphere, with the field-of-view angle Φ_2 and elevation angle θ_2. The linear distance between the receiver and emitter is R, and the common volume of the angle of view and the angle of the beam is V.

When there is an elevation difference between the transmitting and receiving ends, as shown in Fig. 44.7, the line between T_x and R_x is not horizontal because of the height difference between them; at this point, the transmission and reception elevations become the same as θ_1' and θ_2' [41]. Thus, the received power calculation does not refer to R_x, but to a point on a circle centered on T_x, with a radius r and a height equal to T_x. In Fig. 44.7, P_r' is the received power at pointR_x'.

The receiver-power parameter P_r of the receiver R_x is the reference plane in the horizontal plane where the line between T_x and R_x is located. In Fig. 44.7, the

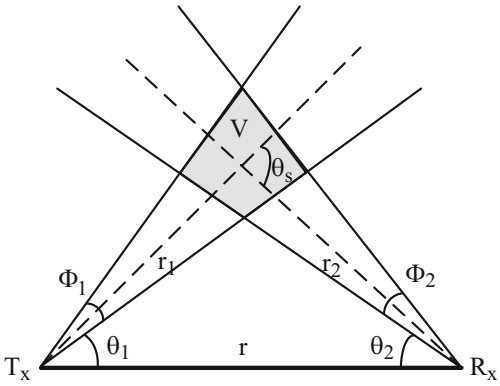


Fig. 44.6 NLOS link model [1]

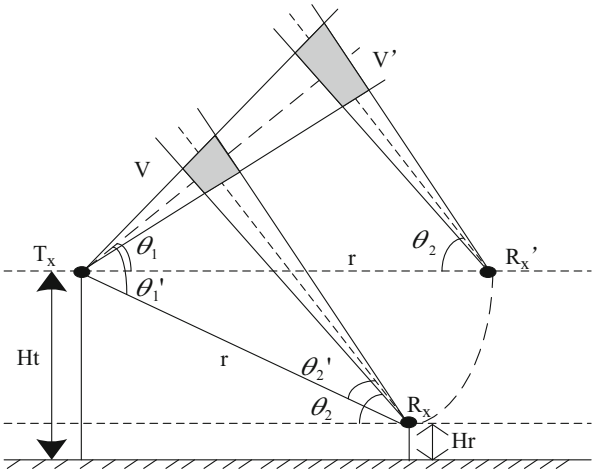


Fig. 44.7 Extremely poor UV NLOS communication [1]

transmit and receive elevations have become the same as θ_1' and θ_2', the changes are recorded as $\Delta\theta_1$ and $\Delta\theta_2$, and the other parameters are kept constant.

He [42] used a 250-nm solar-blind ultraviolet light as the light source in a simulation experiment. They obtained the receiving power after correction for certain heights: $H_t =$ 0.1 km andH_r = 0.04 km, H_t = 0.2 km and H_r = 0.05 km. The measured P_r' is compared with the calculated P_r and the results are shown in Figs. 44.8, 44.9, 44.10, and 44.11. The setting of the extinction coefficient can be found in [43].

As can be seen from the figures, P_r is generally better than P_r'. Therefore, when the ultraviolet communication equipment is networking, each node should look for a node with a small height difference to forward to in the routing process; if there is no available node with a small height difference, the transmission elevation angle should be reduced as much as possible, and the receiving elevation angle should be increased appropriately so that the receiving power is higher.

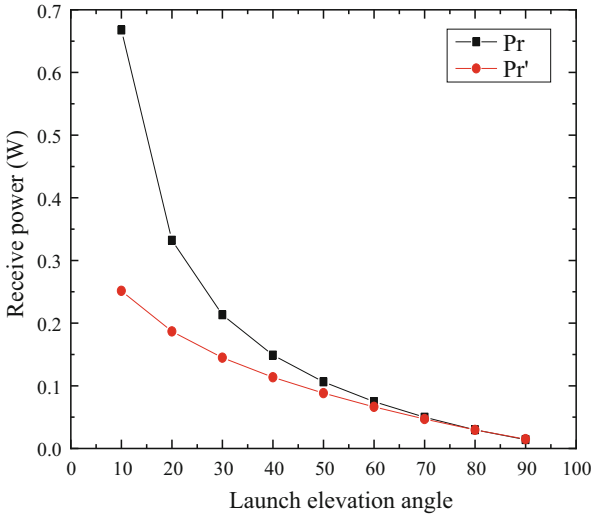


Fig. 44.8 $\theta_2 = 60°$ receiving power [43]. ($H_t = 0.1$ km, $H_r = 0.04$ km)

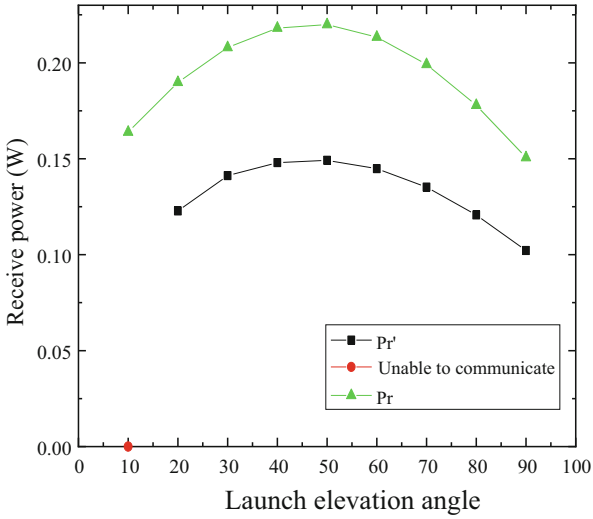


Fig. 44.9 $\theta_1 = 30°$ receiving power [43]. ($H_t = 0.1$ km, $H_r = 0.04$ km)

Pulse-Broadening Effect of NLOS Ultraviolet Communication

Owing to the different scattering points of different photons and different transmission paths after scattering by the common scatterer, the photons reach the receiving end at different times. In addition, multiple scattering may occur in practice, which eventually leads to pulse broadening of the received signal. Solving the pulse-broadening problem plays a key role in determining the decision-threshold level

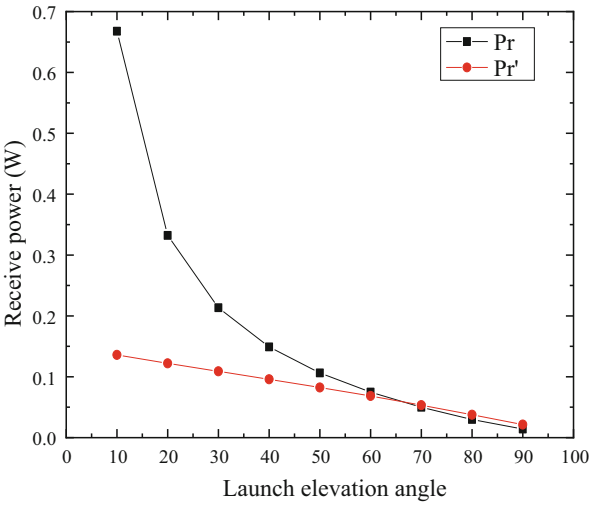


Fig. 44.10 $\theta_2 = 60°$ receiving power [43]. ($H_t = 0.2$ km, $H_r = 0.05$ km)

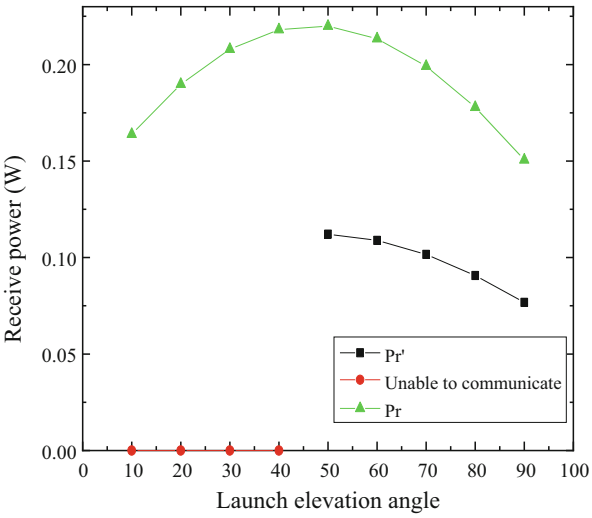


Fig. 44.11 $\theta_1 = 30°$ receiving power [43]. ($H_t = 0.2$ km, $H_r = 0.05$ km)

and decision time to reduce the intersymbol interference (ISI); this improves the system transmission rate and reduces the bit-error rate (BER) [44].

Song simulated the effect of multiple scattering on pulse broadening. The simulation parameters are presented in [40]. Figure 44.12 shows the simulation waveform of the pulse response at the receiving end when the receiving and transmitting elevation angles are the same. The figure shows the degree of pulse broadening after the pulse signal reaches the receiver through the channel. With an increase in

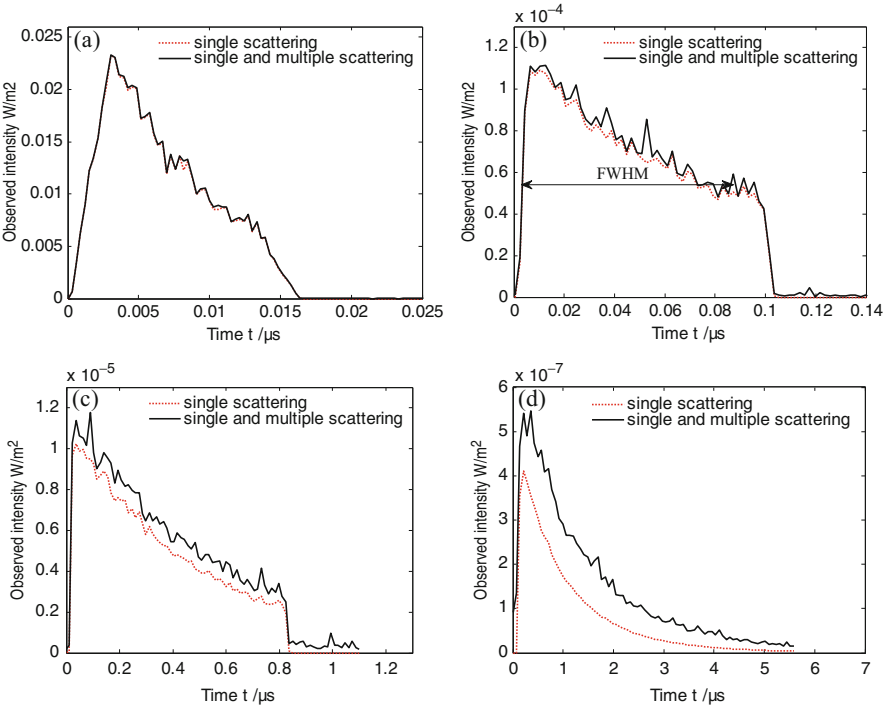


Fig. 44.12 Impulse responses for single and multiple scattering [40]. (**a**) $\theta_t = \theta_r = 10^{\circ}$; (**b**) $\theta_t = \theta_r = 40^{\circ}$; (**c**) $\theta_t = \theta_r = 70^{\circ}$; (**d**) $\theta_t = \theta_r = 90^{\circ}$

the elevation angle of the receiver, the pulse-broadening phenomenon becomes more obvious.

When considering the same transceiver elevation, we should also consider the different transceiver elevations and the pulse-broadening degree contrast. When simulating different elevation angles, the parameters were set the same as above, and the pulse broadening at different elevation angles was compared. The simulation results are shown in Fig. 44.13. A longitudinal comparison of Fig. 44.13a, c, e shows that when the elevation of the receiver is fixed, the larger the elevation of the transmitter, the more serious is the pulse-broadening phenomenon.

Influence of Atmospheric Turbulence on Ultraviolet NLOS Communication

The random fluctuation of local atmospheric physical properties (temperature, pressure, and air velocity) results in random changes in the atmospheric refractive index [45]. In addition, a light beam in the atmosphere will diverge, distort, and drift,

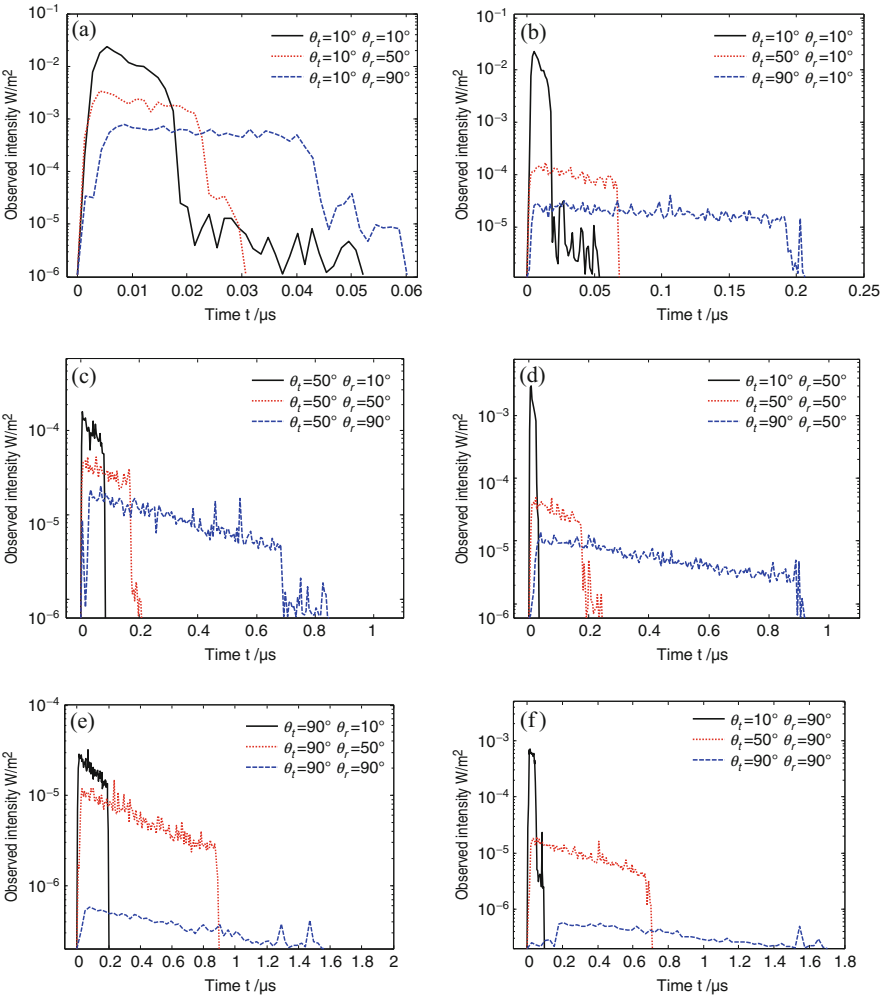


Fig. 44.13 Transmit–receive elevation and pulse broadening [40]. (**a**) $\theta_t = 10^\circ,\ \theta_r = 10^\circ,\ 50^\circ,\ 90^\circ$; (**b**) $\theta_t = 10^\circ,\ 50^\circ,\ 90^\circ,\ \theta_r = 10^\circ$; (**c**) $\theta_t = 50^\circ,\ \theta_r = 10^\circ,\ 50^\circ,\ 90^\circ$; (**d**) $\theta_t = 10^\circ,\ 50^\circ,\ 90^\circ,\ \theta_r = 50^\circ$; (**e**) $\theta_t = 90^\circ,\ \theta_r = 10^\circ,\ 50^\circ,\ 90^\circ$; (**f**) $\theta_t = 10^\circ,\ 50^\circ,\ 90^\circ,\ \theta_r = 90^\circ$

owing to the influence of turbulence, resulting in flickering and jittering in the received light intensity [46].

In weak turbulence, the Rytov solution of the wave equation describes scintillation according to the log-amplitude variance of the received wave. The variance of the log-amplitude fluctuation χ of a plane wave over a length-l channel is [47]

$$\sigma_x^2(L) = 0.56\left(\frac{2\pi}{\lambda}\right)^{7/6}\int_0^L C_n^2(x)(L-x)^{5/6}dx. \tag{44.1}$$

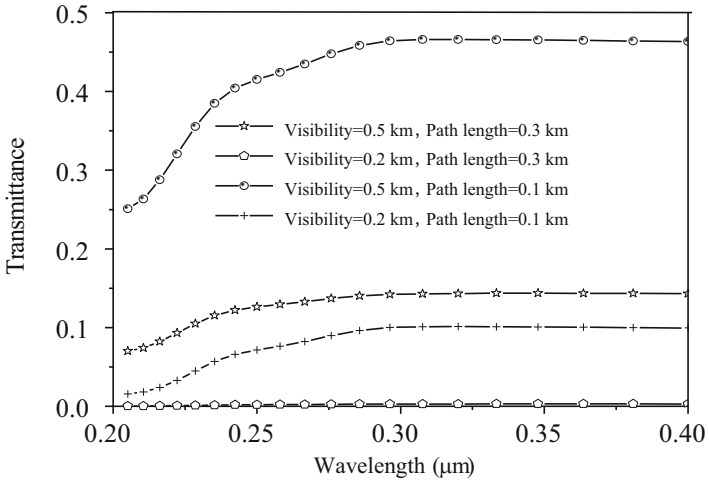


Fig. 44.14 Transmittance of different visibilities and transmission distances [48]

For spherical waves, the variance of the log-amplitude wave χ is [47]

$$\sigma_x^2(L) = 0.56\left(\frac{2\pi}{\lambda}\right)^{7/6}\int_0^L C_n^2(x)\left(\frac{x}{L}\right)^{5/6}(L-x)^{5/6}dx. \tag{44.2}$$

Feng [48] simulated the characteristics of atmospheric ultraviolet transmittance and analyzed it under different transmission distances, visibilities, climates, propagation directions, altitudes, etc.

Figure 44.14 shows the atmospheric transmittance for foggy days with visibilities of 0.2 km and 0.5 km and transmission distances of 0.1 km and 0.3 km. It can be seen that when the communication distance is fixed, the greater the visibility, the better the communication effect. When the visibility is fixed, the lower the transmission rate, the worse the communication quality.

Key Technologies of UV Self-Organizing Networks

An ad hoc network is a wireless mobile communication network composed of a group of mobile nodes; each node is equipped with a wireless transceiver and there is no preset infrastructure. Network communication is established temporarily. Because of its self-organization and distributed nature, it has wide application prospects in the military field. Mobile nodes in a network use their own wireless-transceiver devices for information exchange [1].

Node-Localization Algorithm of an Ultraviolet Optical Communication Network

Ke divided the UV communication network topologies into three types: ring, star, and mesh, as shown in Fig. 44.15a–c.

He discussed commonly used node-location algorithms [43]: trilateration and triangulation. The principle of trilateration is shown in Fig. 44.16. The node coordinates of three different regions are $A(x_1, y_1)$, $B(x_2, y_2)$, and $C(x_3, y_3)$. The distances from the unknown node $D(x, y)$ to the three anchor nodes A, B, and C are, respectively, d_1, d_2, and d_3, which can be obtained from the distance formula between the two points [48].

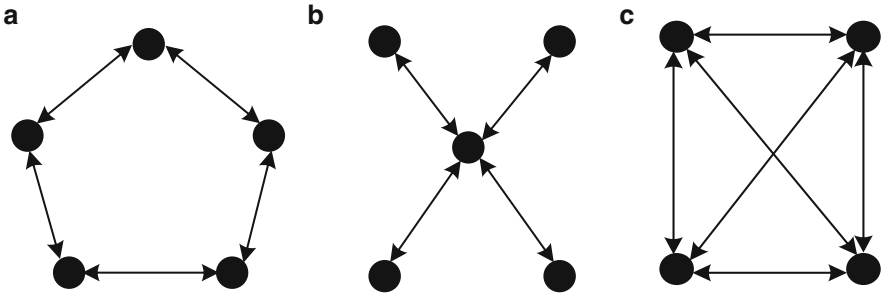


Fig. 44.15 UV network topologies [1]. (**a**) Ring network structure; (**b**) Star network structure; (**c**) Mesh network structure

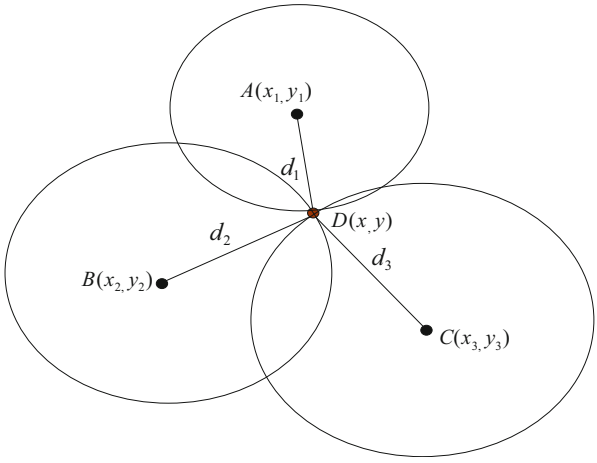


Fig. 44.16 Sketch of the trilateration location algorithm [48]

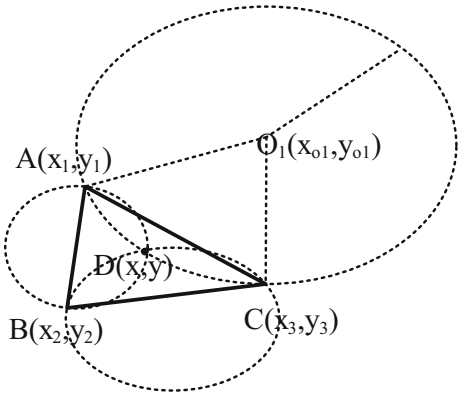


Fig. 44.17 Sketch of the triangulation location algorithm [48]

$$\begin{cases} d_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2} \\ d_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2} \\ d_3 = \sqrt{(x - x_3)^2 + (y - y_3)^2} \end{cases} \tag{44.3}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2(x_2 - x_1) & 2(y_2 - y_1) \\ 2(x_2 - x_3) & 2(y_2 - y_3) \end{bmatrix}^{-1} \begin{bmatrix} x_2^2 - x_1^2 + y_2^2 - y_1^2 + d_1^2 - d_2^2 \\ x_2^2 - x_3^2 + y_2^2 - y_3^2 + d_3^2 - d_2^2 \end{bmatrix} \tag{44.4}$$

The principle of triangulation is shown in Fig. 44.17. The coordinates of known nodes A, B, and C are $A(x_1, y_1)$, $B(x_2, y_2)$, and $C(x_3, y_3)$, and the coordinates of node D are assumed to be (x, y). The coordinates and radius of the center point of a circle are determined using Eq. 44.5.

$$\begin{cases} \sqrt{(x_{O1} - x_1) + (y_{O1} - y_1)} = r_1 \\ \sqrt{(x_{O1} - x_3) + (y_{O1} - y_3)} = r_1 \\ (x_1 - x_2)^2 + (y_1 - y_3)^2 = 2{r_1}^2 - 2{r_1}^2 \cos\alpha \end{cases} \tag{44.5}$$

Ultraviolet Self-Organizing-Network Multiuser Detection Technology

In [49], Liang analyzed the throughput of ad hoc networks based on multiuser detection technology and proposed the principles that should be followed in UV networks. Based on this, an adaptive recursive least-squares (RLS) algorithm was

introduced into optical code-division multiple access (OCDMA), and an improved variable forgetting-factor RLS algorithm was proposed [50].

The RLS adaptive algorithm differs from conventional least-squares in that it is an exponentially weighted least-squares algorithm. Using the weighted sum of squares of errors as the cost function and the exponential weighted sum instead of the constrained minimum output energy, $CMOE(\mathbf{c}_1) = E\left\{\left(\boldsymbol{c}_1^T\boldsymbol{r}\right)^2\right\}$, the following formula [50] is obtained:

$$\begin{cases} \min \sum_{i=1}^{n} \lambda^{n-i}\left[\mathbf{c}_1^T(n)\mathbf{r}(i)\right]^2 \\ \boldsymbol{s.t.}\ \ \boldsymbol{s}_1^T\boldsymbol{c}_1(n) = 1 \end{cases}, \tag{44.6}$$

where λ is the forgetting factor. By using the Lagrange multiplier method, the closed-form solution of the optimization problem can be obtained as follows:

$$\mathbf{c}_1(n) = \frac{\mathbf{R}^{-1}(n)\mathbf{s}_1}{\mathbf{s}_1^T\mathbf{R}^{-1}(n)\mathbf{s}_1}, \tag{44.7}$$

where $\mathbf{R}(n) = \sum_{i=1}^{n} \lambda^{n-1}\left[\mathbf{r}(i)\mathbf{r}^T(i)\right]^2$ represents the autocorrelation matrix of the observed signal.

Assuming the simulation condition of an additive white-Gaussian-noise synchronous OCDMA system, Liang simulated an adaptive multiuser detector using the RLS algorithm. The BER performance curves are shown in Figs. 44.18, 44.19, and

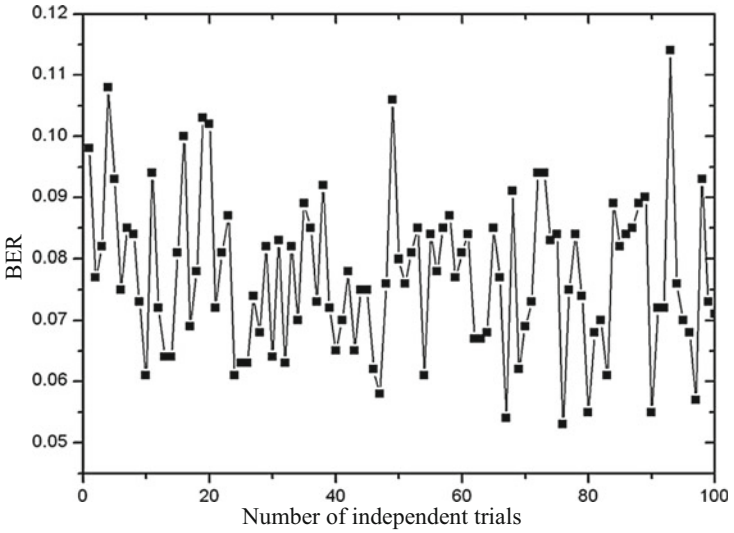


Fig. 44.18 $\lambda = 0.99$. RLS algorithm bit-error rate [49]

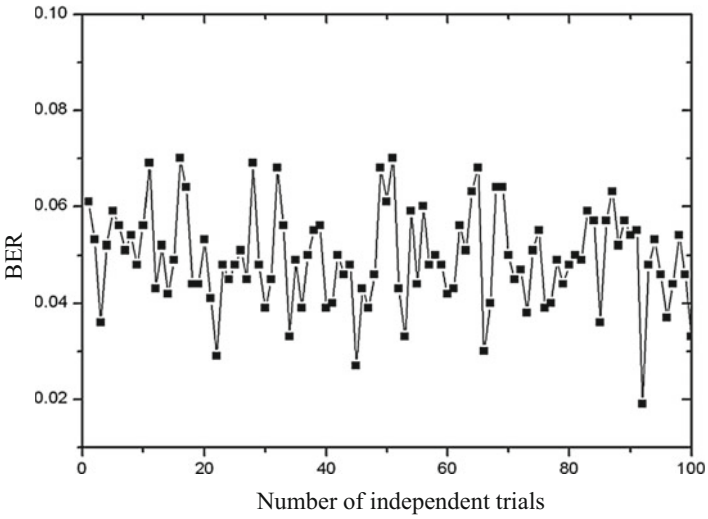


Fig. 44.19 $\lambda = 0.997$. RLS algorithm bit-error rate [49]

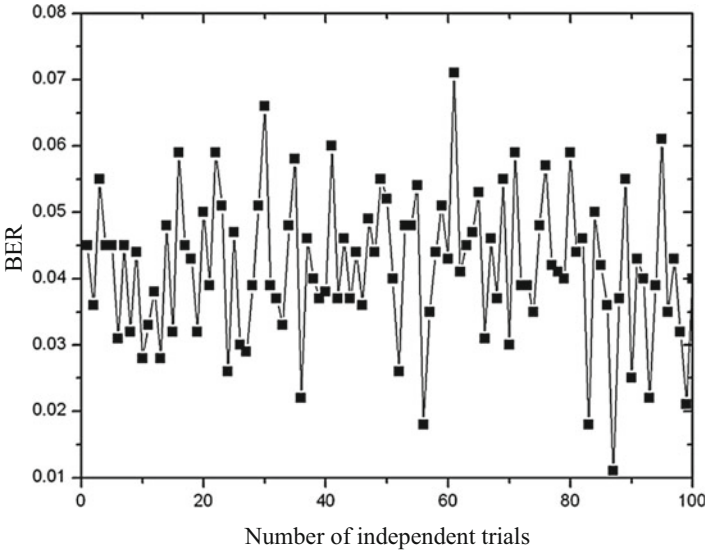


Fig. 44.20 $\lambda = 0.9997$. RLS algorithm bit-error rate [49]

44.20. The graph shows that when conducting blind multiuser OCDMA detection using the RLS algorithm, the expected user signal is tracked relatively well. The estimated bit-error rate of the information characters sent by the expected user is lower; the blind multiuser OCDMA detection can be completed very well.

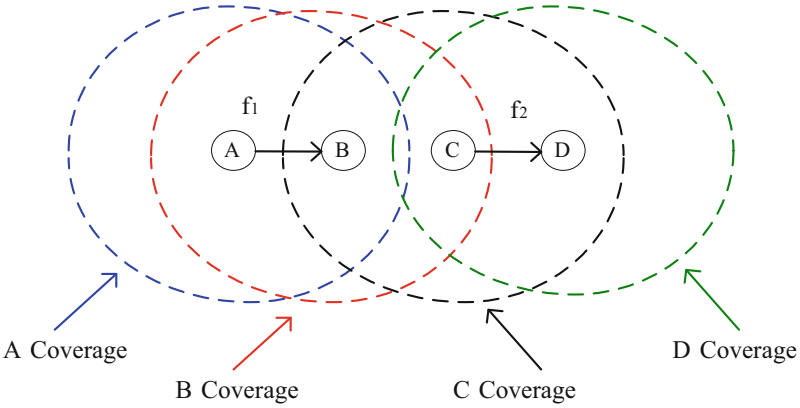


Fig. 44.21 Node-chain topology (exposed terminal) [51]

MAC-Layer Protocol

In the protocol stack of the wireless ad hoc network, the MAC protocol plays an important role. It is the manager that defines the medium in which the data are sent and received over a wireless channel. In [51], Chen analyzed the performance of ad hoc networks with MAC-layer saturation using the IEEE 802.11 distributed-coordination function (DCF) protocol by Markov modeling. In this chapter, history-based dynamic fair backoff (HDFB) and the new DCF (NDCF) improved backoff algorithms are proposed to solve different fairness problems.

For the above-mentioned improved algorithm, Chen et al. [51] carried out a simulation based on IEEE 802.11 DCF, in which the ultraviolet communication single-scattering link model was adopted to replace the default two-ray ground-reflection model in the NS2 network-simulation software. A typical four-node chain in the ad hoc network topology was selected for simulation, as shown in Fig. 44.21.

The simulation results are shown in Figs. 44.22 and 44.23 [52], which show that for the HDFB algorithm, the throughput difference between node *C* and node *A* is approximately 20 kbit/s, with the best fairness. For the NDCF algorithm, the throughput difference between nodes *C* and *A* was approximately 290 kbit/s, which is fair. For the binary exponential backoff (BEB) algorithm, the throughput difference between nodes *C* and *A* was up to 536 kbit/s, and the fairness is the worst.

UV Ad hoc: Network Routing Protocol

Ultraviolet (UV) communication is characterized by its small transmitter and receiver, low cost, indirect communication, high security, very low background noise, and potential for high-speed communication. This makes it an ideal complement to traditional communication. Requirements for ad hoc–network routing

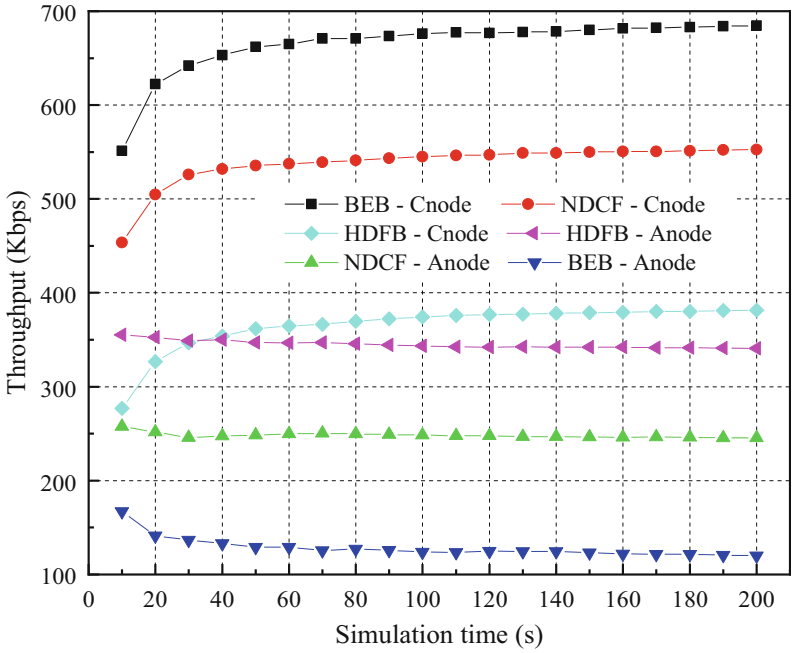


Fig. 44.22 Node-chain topology throughput [52]

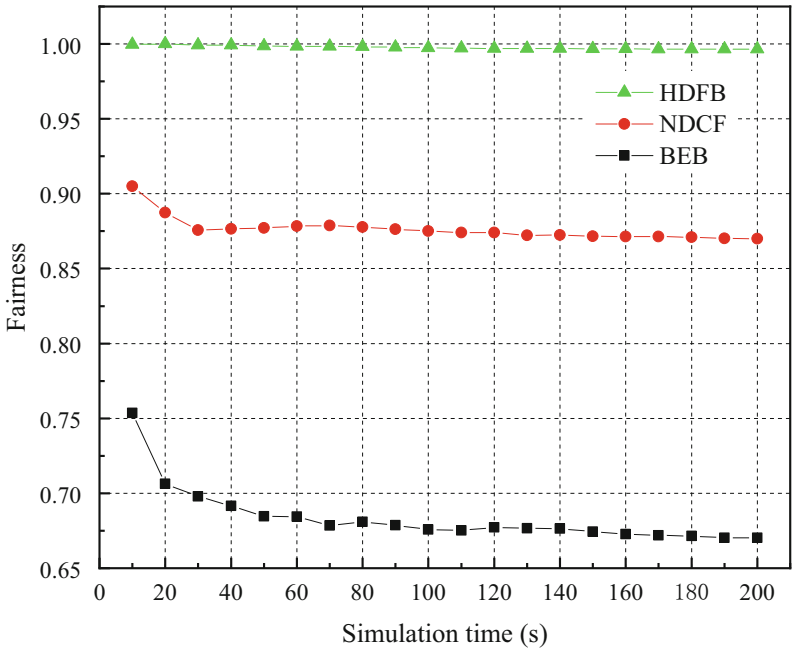


Fig. 44.23 Node-chain topology fairness index [52]

protocols, such as fast convergence, loop-free routing, infinite-computation avoidance, low overhead in control management, and support for one-way channels, have been presented [53].

Yang [53] proposed a position- and velocity-aided routing protocol (PVAR) using geographic-location information and nodes to maintain the effectiveness of the routing as long as possible, and effectively reduce the routing overhead in the network.

In the simulation, Yang set the number of nodes in the network to 50 and 100. The nodes in the network were limited to a 670 × 670 m square region. The velocities of the nodes were 1–10 ms, 10–20 ms, and 20–30 ms, and followed a Gaussian distribution. The PVAR and ad-hoc on-demand distance vector (AODV) protocols were then compared.

The simulation results for the packet arrival rate and routing overhead are shown in Figs. 44.24, 44.25, 44.26, and 44.27, respectively. From the figure, we can see that the performance of the PVAR protocol is better than that of the AODV protocol. At different speeds, the packet arrival rate of PVAR was approximately 90%, whereas the packet arrival rate of AODV was less than 90% when the node speed was between 20 and 30 m/s. The performance of PVAR was better than that of AODV, and PVAR maintained a low routing overhead. However, when the node speed in the network was between 20 and 30 m/s, the performances of PVAR and AODV degraded rapidly, and their jitter was very large.

To find the optimal path from the source node to the destination node, Wu used unidirectional-link routing based on an ant-colony algorithm for UV wireless mesh network (ULR-ACA-UVWMN); that is, in an ultraviolet wireless mesh-network topology using one-way and two-way link methods, the ant-colony algorithm and

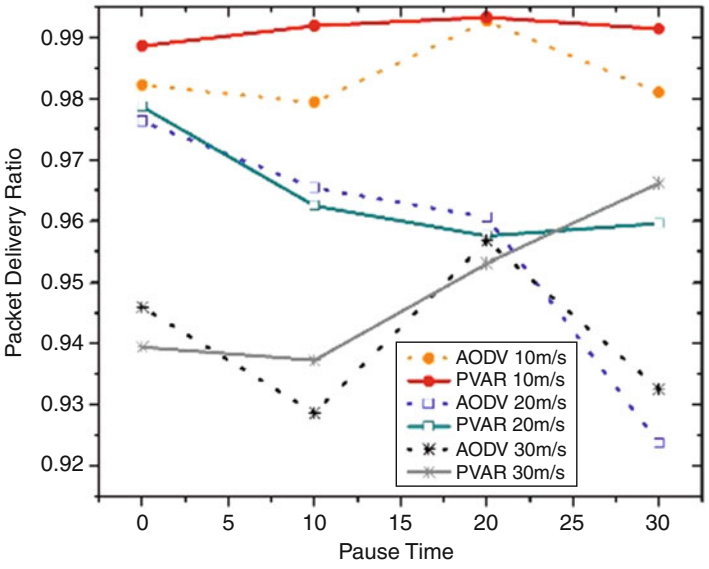


Fig. 44.24 50-node network packet arrival rate [53]

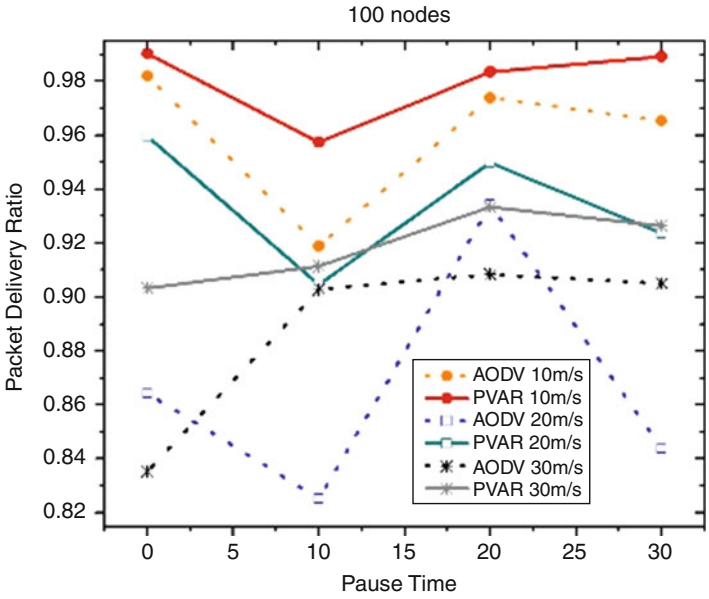


Fig. 44.25 100-node network packet arrival rate [53]

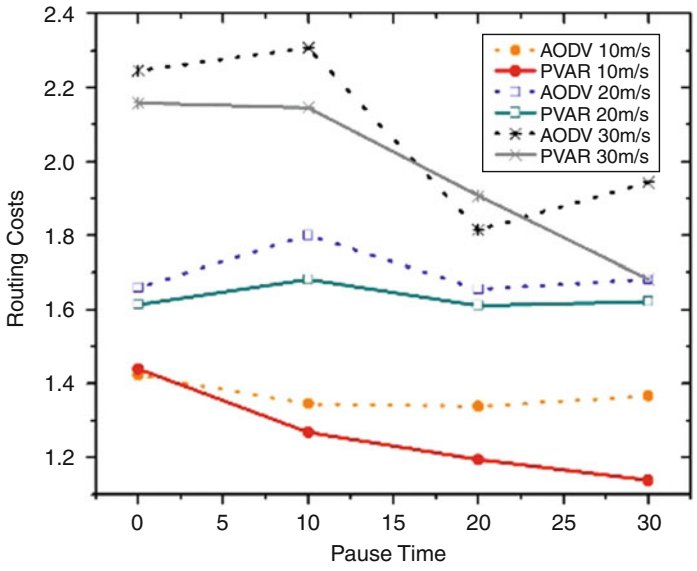


Fig. 44.26 50-node network routing overhead [53]

the ULR-ACA-UVWMN algorithm, which only supports two-way links, were simulated [54].

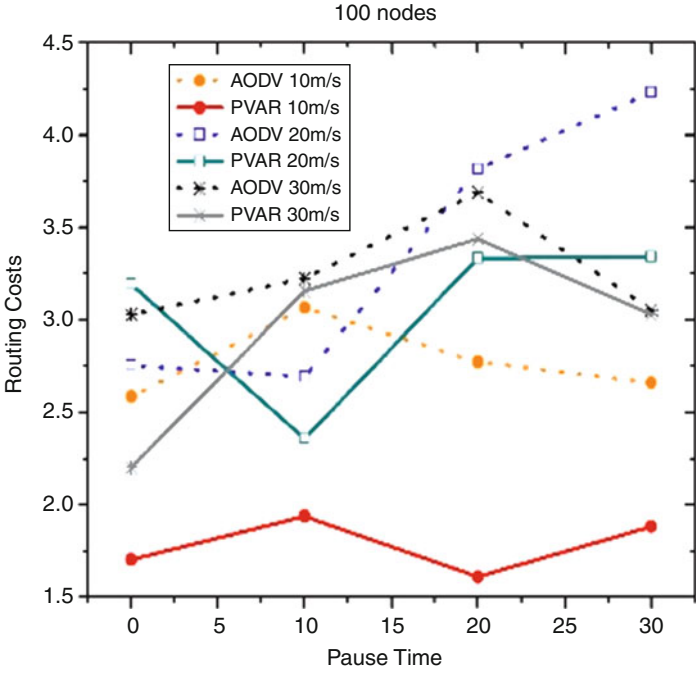


Fig. 44.27 100-node network routing overhead [53]

The simulation results are shown in Figs. 44.28 and 44.29. From the graph, it can be seen that the convergence speed of the ant-colony algorithm, which only supports bidirectional links, to reach the optimal path, is lower than that of the ULR-ACA-UVWMN algorithm.

Analysis of Energy Saving in UV Wireless Sensor Networks

Although multi-hop communication using ultraviolet light can save energy, the greater the number of hops, the more the routing load and MAC-layer load will increase, which is not conducive to energy saving; therefore, there must be an optimal hop number for multi-hop communication to optimize the network performance [55].

Using Fig. 44.30 as the simulation model, 11 nodes are configured in a 600 m × 600 m region, with an initial energy of 10 J for each node. Node 1 is the source node, and node 11 is the sink node. $d_{\max} = 550$ m and $d_{\min} = 55$ m, which means that data can be relayed from the source node to the sink node over a distance of up to 10 hops. The simulation compared and analyzed the energy consumption of the three communication modes, as shown in Fig. 44.31. As can be seen from the figure, five-hop communication is better than one-hop and nine-hop communication, in terms of

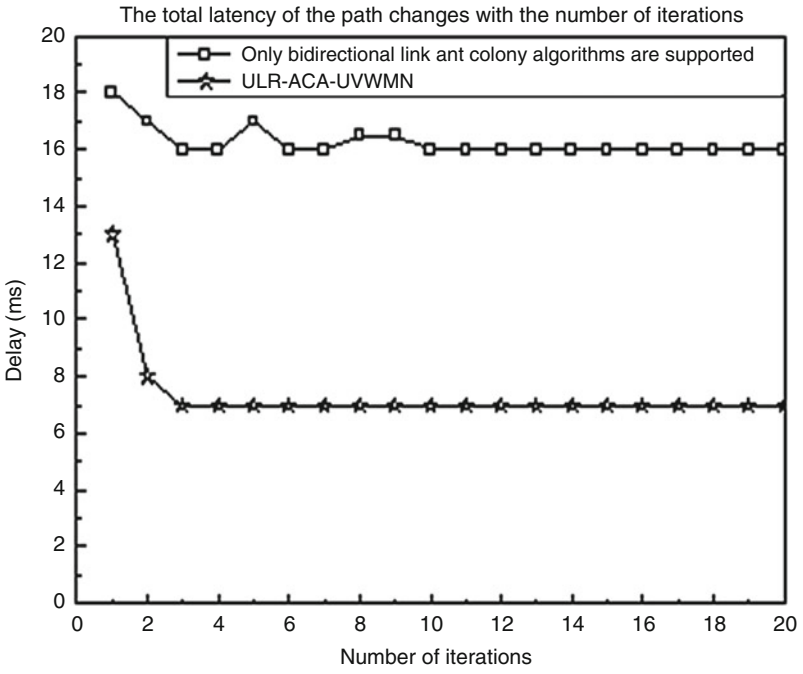


Fig. 44.28 Four-node network topology [54]

saving energy. The number of hops should also be combined with the actual application, not simply maximized.

Summary

Based on the development of ultraviolet communication at home and abroad, single- and multiple-scattering models for ultraviolet communication were introduced and the different links in ultraviolet communication were analyzed. The effects of pulse broadening and atmospheric turbulence on ultraviolet communication in NLOS links were discussed, and key technologies, such as node-localization algorithms, multi-user detection, MAC-protocol fairness, and routing protocols, were summarized. Currently, ultraviolet optical communication can be developed in the following ways.

1. The atmosphere strongly absorbs ultraviolet light, which attenuates to different degrees in the transmission channel, resulting in poor communication quality. Network relays can be used to solve this problem by increasing the number of appropriate nodes. In addition, full-duplex communication technology is necessary to realize practical ultraviolet communication. Because of UV's short wavelength and large bandwidth, the transmission rate of the optical signal can be

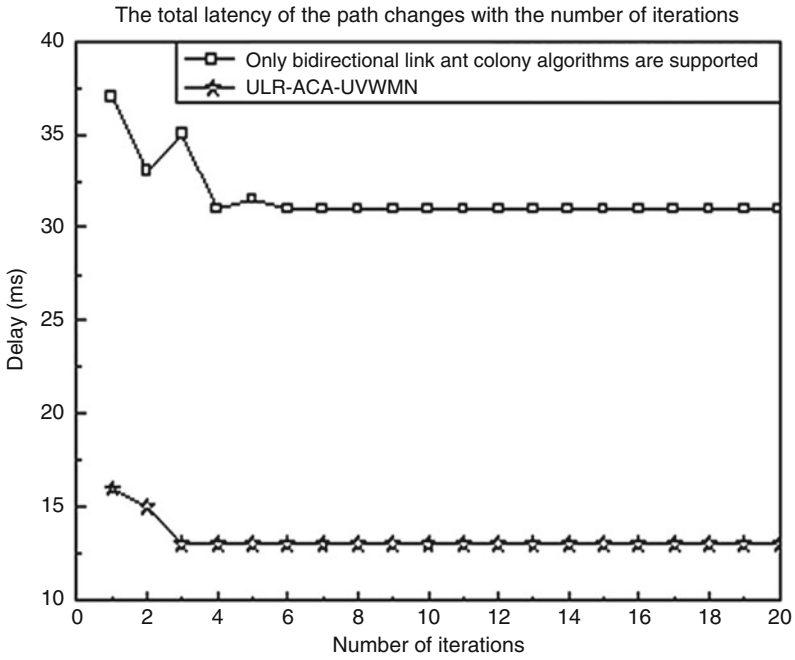


Fig. 44.29 Eight-node network topology [54]

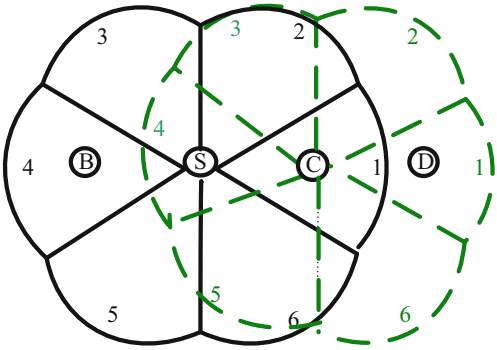


Fig. 44.30 UV sensor-node model [55]

effectively increased using space-division multiplexing; thus, long-distance coverage of UV communication can be achieved.

2. The key and difficult point of studying a UV sensor network is multiuser access in an uplink. Interference between adjacent nodes is consistently a major problem. At present, the interference between adjacent nodes can be reduced by space-division multiplexing.
3. UV NLOS positioning technology is the key direction for integrating UV NLOS communication and positioning in the future, which can be used to help stabilize

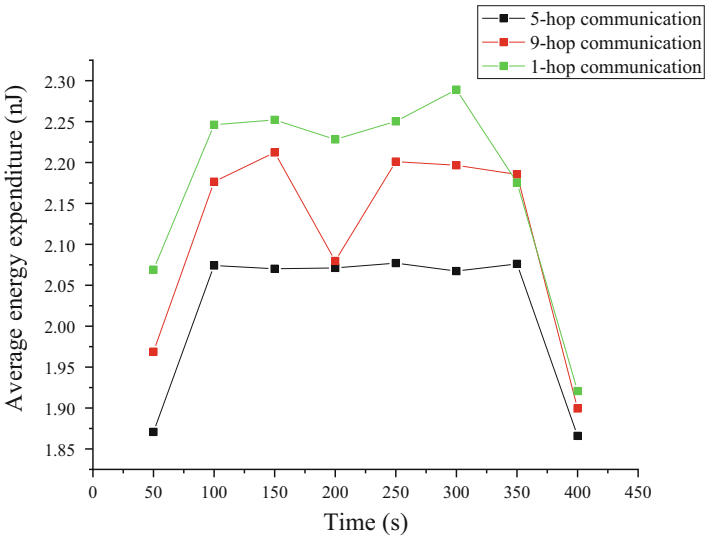


Fig. 44.31 Comparison of average energy consumption for 1-hop, 9-hop, and 5-hop communications [56]

the communication. In the future, the main challenge of UV NLOS positioning technology will be to deduce the distance and azimuth of the transmitter using the information in the receiving power.

References

1. X Ke. Theory of ultraviolet self-organizing network [M]. Beijing: Science Press, 2011: 40–42.
2. Ding Y, Fan J, Tong S, et al. Range extension method for communication distance of atmospheric channel ultraviolet optical communication system [J]. Acta Photonica Sinica, 2012, 41(9): 1047–1052.
3. Sunstein D E. A scatter communications link at ultraviolet frequencies [D]. Cambridge: Massachusetts Institute of Technology, 1968: 35–45.
4. Reilly D M. Atmospheric optical communications in the middle ultraviolet [D]. Cambridge: Massachusetts Institute of Technology, 1976: 77–90.
5. Geller M, Keenan T E, Altman D E, et al. Optical non-line-of-sight covert, secure high data communication system [P]. U.S. Patent: 4, 493, 114. 1985: 1–8.
6. Luettgen M R, Shapiro J H, Reilly D M. Non-line-of-sight single-scatter propagation model [J]. JOSA A, 1991, 8(12): 1964–1972.
7. Yuan R, Ma J. Review of ultraviolet non-line-of-sight communication [J]. China Communications, 2016, 13(6): 63–75.
8. Shaw G A, Siegel A M, Model J, et al. Recent progress in short-range ultraviolet communication [J]. Unattended Ground Sensor Technologies and Applications VII, 2005, 1(5796): 214–225.
9. Chen G, Abou-Galala F, Xu Z, et al. Experimental evaluation of LED-based solar blind NLOS communication links [J]. Optics express, 2008, 16(19): 15059–15068.

10. Li Y, Wang L, Xu Z, et al. Neighbor discovery for ultraviolet ad hoc networks [J]. IEEE Journal on Selected Areas in Communications, 2011, 29(10): 2002–2011.
11. Wang P, Xu Z. Characteristics of ultraviolet scattering and turbulent channels [J]. Optics letters, 2013, 38(15): 2773–2775.
12. El-Shimy M A, Hranilovic S. Spatial-diversity imaging receivers for non-line-of-sight solar-blind UV communications [J]. Journal of Lightwave Technology, 2015, 33(11): 2246–2255.
13. Ardakani M H, Heidarpour A R, Uysal M. Performance analysis of MIMO NLOS UV communications over atmospheric turbulence channels [C]//2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW). IEEE, 2016: 1–5.
14. Refaai A, Abaza M, El-Mahallawy M S, et al. Performance analysis of multiple NLOS UV communication cooperative relays over turbulent channels [J]. Optics express, 2018, 26(16): 19972–19985.
15. Arya S, Chung Y H. Amplify-and-forward multihop non-line-of-sight ultraviolet communication in the gamma–gamma fading channel [J]. Journal of Optical Communications and Networking, 2019, 11(8): 422–449.
16. Hasan Hariq S, Odabasioglu N. Spatial diversity techniques for non-line-of-sight ultraviolet communication systems over atmospheric turbulence channels [J]. IET Optoelectronics, 2020, 14(5): 327–349.
17. Bhatia V, Jain S, Garg K, et al. Performance analysis of RKHS based detectors for nonlinear NLOS ultraviolet communications [J]. IEEE Transactions on Vehicular Technology, 2021, 70(4): 3625–3639.
18. Arslan C H, Dagefu F T, Moore T J, et al. Optimization of ultraviolet communication links based on finite difference stochastic approximation [J]. Optics Express, 2022, 30(20): 36283–36296.
19. Ni G, Zhong S, Liu L, et al. Research on free atmosphere ultraviolet optical communication [J]. Optical technology, 2000, 26(4): 297–303.
20. Lan T, Ni G. Simulation study on atmospheric transmission characteristics of ultraviolet communication [J]. Journal of Beijing Institute of Technology, 2003, 27(4): 419–423.
21. Lan Y, Xiao S, Luo Y, et al. Research on transmission and modulation technology of solar blind ultraviolet optical communication system [J]. Optoelectronic technology, 2007, 25(03): 206–211.
22. Zhao M, Xiao S, Wang X, et al. Research on ultraviolet optical communication system based on LED [J]. Advances in lasers and optoelectronics, 2010, 47(4): 19–24.
23. Ke X, He H, Chen X. A New MAC Layer Avoidance Algorithm for Ultraviolet Ad Hoc Communication Networks [J]. Photoelectron. laser 2010, 21(7): 1002–1006.
24. Ke X, Chen Ji, Hou Z. Directional Media Access Control Protocol for Ultraviolet Non-line-of-Sight Communication [J]. Photoelectron. laser, 2011, 22(8): 1190–1195.
25. Zhao T, Ke X. Monte Carlo Method for Simulation of Non-Direct UV Light Scattering Coverage [J]. Acta Physica Sinica, 2012, 61(11): 285–296.
26. Wu C, Zuo Y, Fan C, et al. Research on MAC layer power control in ultraviolet optical communication network [J]. Optical communication research, 2013, 39(6): 28–38.
27. Xu C, Zhang H. Monte-Carlo based modeling for ultraviolet non-line-of-sight communication channels with typical obstacles [C]//2015 International Conference on Wireless Communications & Signal Processing (WCSP). IEEE, 2015: 1–5.
28. Yang G, Li X, Chen M, et al. A New Time-Division Multiple Access Neighbor Discovery Algorithm for Ultraviolet Ad Hoc Networks [J]. Photoelectron. laser, 2015, 26(6): 1074–1080.
29. Song P, Song F, Li Y, et al. Multi-user interference problem of non-direct UV optical communication networking [J]. Acta Photonica Sinica, 2016, 1(9): 159–166.
30. Song P, Zhou X. Design and Communication Performance Analysis of Ultraviolet Mobile Ad Hoc Network Nodes [J]. Acta Optica Sinica, 2018, 38(3): 290–297.
31. Wang P, Wang J, Li J, et al. Experiment on UV light communication multiplex modulation system based on low-voltage mercury lamp [J]. Optical communication technology, 2019, 43(8): 21–24.
32. Zhang J, Zhou M. Design of a low-latency ultraviolet optical communication system [J]. Laser Magazine, 2020, 41(4): 150–154.

33. Cao T, Gao X, Wu T, et al. Single-scatter path loss model of LED-based non-line-of-sight ultraviolet communications [J]. Optics letters, 2021, 46(16): 4013–4016.
34. Cao T, Wu T, Pan C, et al. Single-collision-induced path loss model of reflection-assisted non-line-of-sight ultraviolet communications [J]. Optics Express, 2022, 30(9): 15227–15250.
35. Chne G, Xu Z, Ding H, et al. Path loss modeling and performance trade-off study for short-range non-line-of-sight ultraviolet communications [J]. Optics express, 2009, 17(5): 3929–3940.
36. Zhao T, Feng Y, Ke X, et al. Research on node coverage in "solar blindness" ultraviolet optical communication network. Acta Optica Sinica, 2010, 30(8): 2229–2248.
37. Feng Y. Research on atmospheric channel of ultraviolet optical communication [D]. Xi'an: Xi'an University of Technology, 2011: 20–48.
38. Drost R J, Moore T J, Sadler B M. UV communications channel modeling incorporating multiple scattering interactions [J]. JOSA A, 2011, 28(4): 686–695.
39. Song P, Su C, Zhao T, et al. Capacity Analysis of Non-direct Vision Non-coplanar UV Optical Communication Channel Based on Traversal Microunit Method [J]. Journal on Communications, 2019, 40(05): 144–152.
40. Song P. Research on channel transmission characteristics of wireless ultraviolet mobile ad hoc network [D]. Xi'an: Xi'an University of Technology, 2017: 35–43.
41. He H, Ke X. Mie scattering mechanism in ultraviolet optical communication [J]. Journal of Applied Sciences, 2012, 30(03): 245–250.
42. He H, Ke X, Zhao T. Research on NLOS Single Scatter Link Model Based on Altitude [J]. Laser technology, 2011, 35(04): 495–498.
43. He H. Research on key technologies of wireless ultraviolet optical communication and its networking [D]. Xi'an: Xi'an University of Technology, 2011: 37–45.
44. Song P, Ke X, Xiong Y, et al. Pulse broadening effect of non-direct ultraviolet light in non-coplanar communication system [J]. Acta Optica Sinica, 2016, 36(11):55–64.
45. Aung T Y, Arya S, Chung Y H. Performance dependence of non-line-of-sight ultraviolet communications on atmospheric parameters of the ultraviolet channel [J]. Optics Communications, 2019, 443(1): 7–11.
46. Biswal M R, Arya S, Chung Y H. Effect of turbulence and noise on ultraviolet and mid-infrared spectrum in optical wireless communications [J]. Photonic Network Communications, 2020, 39(3): 181–186.
47. Zuraski S M, Beecher E, McCrae J E, et al. Turbulence profiling using pupil plane wavefront data derived Fried parameter values for a dynamically ranged Rayleigh beacon [J]. Optical Engineering, 2020, 59(8): 1–14.
48. He H, Ke X, Feng Y. Research on localization in wireless "solar blind" ultraviolet light grid network [J]. Laser technology, 2010, 34(05): 607–610.
49. Liang Wei. Research on multi-user detection technology in ultraviolet ad hoc network [D]. Xi'an: Xi'an University of Technology, 2011: 37–45.
50. Wu P, Ke X, Liang W. An Improved Blind Multi-User Detection Algorithm in Wireless Optical Code Division Multiple Access [J]. Acta Optica Sinica, 2012, 32(2): 86–91.
51. Ke X, He H, Chen X. Dynamic fairness avoidance algorithm based on history in ultraviolet ad hoc network. Laser Magazine, 2010, 31(3):32–46.
52. Chen X. Research on MAC layer protocol fairness of ultraviolet ad hoc network [D]. Xi'an: Xi'an University of Technology, 2010: 35–43.
53. Yang P Research on routing protocol of ultraviolet ad hoc network [D]. Xi'an: Xi'an University of Technology, 2011: 42–51.
54. Wu C. Research on routing protocol of ultraviolet optical communication network based on ant colony algorithm [D]. Xi'an: Xi'an University of Technology, 2011: 43–46.
55. Ke X, Chen J. Research and simulation of energy saving of ultraviolet wireless sensor network [J]. Laser technology, 2013, 37(2): 251–255.
56. Deng L. Research on routing protocol of ultraviolet wireless sensor network [D]. Xi'an: Xi'an University of Technology, 2011: 30–45.

Experimental Study on Ultraviolet Optical Communication Noise

45

Contents

Introduction 1648
Foreign and Domestic Research Progress 1650
Foreign Research Status 1650
Domestic Research Status 1653
Theoretical Model of Ultraviolet Background Light 1655
Theoretical Model of Solar Radiation to the Earth 1655
UV Optical Communication Path Loss Model 1663
UV Light Scattering Channel Characteristics Under Different Weather Conditions 1674
Experimental Study of the Noise Model Under Different Weather Conditions 1684
Experimental Setup 1684
Near Earth Background Light Noise Model 1685
Analysis of Light Intensity Distribution Characteristics of Ultraviolet Optical Communication 1692
Summary 1698
References 1699

Abstract

Ultraviolet (UV) wireless communication is a new method that uses 200 nm–280 nm "solar-blind"-band ultraviolet light as a carrier. It can realize nonline-of-sight (NLOS) communication in complex environments through the scattering of atmospheric molecules and aerosol particles. Under different weather conditions, the bit error rate and path loss of a UV optical communication system change with the particle concentration, particle size, and geometric parameters of the transmitter and receiver in the atmosphere. At different altitudes and times, the bit error rate and path loss of UV optical communication systems change with the concentration of atmospheric molecules and amount of background noise. In this study, the channel characteristics of the UV light scattering intensity distribution and path loss of different rainfall intensities, fog concentrations, and sand and dust types were studied using single-scattering and multiple-scattering UV communication link transmission models. The specific research contents are as follows:

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_45

1. To study the communication performance of NLOS ultraviolet optical communication in low-visibility environments, such as rain, fog, and dusty weather, a new empirical optical attenuation model was used for the ultraviolet atmospheric-channel attenuation analysis, the calculation of multiple-scattering transmission models, and the short distance of UV optical communication in low-visibility environments. Moreover, the asymmetric factors in the scattering phase function were simplified. The physical characteristics of aerosol particles under several weather conditions were analyzed; an empirical model of ultraviolet light attenuation in low-visibility weather was established, based on Mie scattering theory; and the ultraviolet light attenuation parameters of different aerosols at different concentrations were obtained.
2. Changes in ultraviolet background radiation at different times and seasons under sunny weather conditions were derived, and the effectiveness of the theoretical model was verified by comparing it with measurement data. The differences between the measured data and the theoretical models in cloudy and hazy weather were compared. Based on a theoretical model of the surface, the ultraviolet light noise distribution at high altitudes was measured, and the change in high-altitude noise was restored. The experimental results showed that with each 1000-m increase in altitude, the daily variation trend of ultraviolet noise remained unchanged, and the ultraviolet radiation value increased by approximately 20–23%. The establishment of this model lays a theoretical foundation and provides an experimental basis for space ultraviolet optical communication.

Keywords

Ultraviolet optical communication · Atmospheric turbulence · Path loss model · High-altitude noise model

Introduction

Because of its strong scattering properties in the atmosphere, the ultraviolet (UV) band is ideal for nonline-of-sight (NLOS) optical communications. The optical transmission process is affected by Rayleigh and Mie scattering, and with an increase in transmission distance, the ultraviolet light intensity is rapidly attenuated; thus, the ultraviolet light signal can only be received within a certain range, which can prevent eavesdropping. This unique characteristic makes it highly valued by the military.

The ultraviolet band can be divided into multiple parts. The solar radiation in the 200-nm to 260-nm band is absorbed by the atmospheric ozone layer about 20 km above the earth's surface; therefore, the radiation in this band on the surface is much lower than the radiation of visible light and near-ultraviolet light. It is often called the ultraviolet "solar-blind region." If the solar-blind band with low

background noise is used as the communication spectrum, the system will be highly sensitive. The receiving end only needs to receive a small amount of ultraviolet light signal to be able to communicate, and a high signal-to-noise ratio can be obtained.

The main advantages of ultraviolet light communication are as follows:

1. Indirect vision: Conventional space laser communication uses direct visual transmissions. There can be no obstacles in the communication channel. Ultraviolet light signals are strongly scattered by atmospheric particles in the atmospheric channel; thus, an obstacle can be bypassed to complete the communication. Nonline-of-sight is a prominent advantage of UV optical communication over other atmospheric laser communications.
2. Low azimuth resolution: First, because ultraviolet light is invisible, it is difficult for the human eye to determine the orientation of the transmitter. Second, because the ultraviolet light signal is scattered in the air after the transmitter emits it and propagates in all directions, it is difficult to determine the specific orientation of the transmitter from the ultraviolet light signal.
3. Low eavesdropping rate: The ultraviolet light signal is strongly absorbed by the atmosphere during atmospheric propagation; therefore, the energy of the ultraviolet light signal is rapidly attenuated, and communication is only effective within a certain range. After exceeding the communication distance, the ultraviolet signal light energy is obliterated by atmospheric background noise. Even with highly sensitive detectors, it is difficult to detect, which ensures a low eavesdropping rate.
4. High interference resistance: The intensity of solar radiation on the near Earth's surface in the ultraviolet band of mountains is very small because it is scattered by the atmosphere. Ultraviolet solar radiation on the near Earth's surface is evenly distributed; this radiation is reflected in the detector side as DC signals; thus, low-pass filters can be easily used to filter out this background noise.
5. Nondirectional: The principle of ultraviolet communication mentioned above makes great use of the atmospheric scattering effect on ultraviolet light signals in ultraviolet communication.

Ultraviolet light signals are scattered in the atmosphere, and signals can be detected within the effective communication distance centered on the light source. It does not require strong directivity or beam calibration devices like conventional space laser optical communication does.

In an ultraviolet system, owing to the very weak received signal and low amplitude, noise can greatly interfere with the signal. If the noise characteristics are not analyzed, the level judgment will be affected, resulting in a serious decrease in the system bit error performance. To reduce the influence of background noise on communication quality and achieve more stable high-altitude ultraviolet communication and a wider communication bandwidth, it is necessary to study the background noise of ultraviolet light in the solar-blind zone.

In this study, the background noise of 220 nm ±10 nm in the boundary band of the solar-blind zone and 300 nm ±10 nm in the near-solar-blind zone is selected for theoretical modeling. The background noise of the band near the ground and at high altitudes is measured. The measurement results are compared with the theoretical model to verify its correctness and obtain a more complete background ultraviolet daily variation model.

When studying an ultraviolet optical communication system, the noise of the system was modeled, a theoretical model was established, and the parameters of the noise model were adjusted to make the noise model consistent with the measured model. The communication performance before and after modeling was compared to improve the communication quality of the ultraviolet optical communication system under rain, fog, snow, and dusty weather conditions.

Foreign and Domestic Research Progress

In recent years, with the rapid development of ultraviolet detection technology, ultraviolet optical communication has gradually become a research hotspot because of its inherent advantages, such as low solar background noise; indirect vision; strong confidentiality; no acquisition, pointing, and tracking (APT) tracking; and good concealment. Owing to its strong scattering properties in the atmosphere, the ultraviolet band is ideal for nonline-of-sight (NLOS) communication.

Foreign Research Status

As early as 1939, the United States discussed the technology and performance of ultraviolet light sources, detectors, and filters for naval communications but did not conduct actual research [1].

In 1968, Sunstein et al. [2] built the first ultraviolet optical communication link, discussed the selection and limitations of devices in detail, and provided experimental results for the system under different parameters. In 1976, under the American Standard Atmospheric Model, Reilly proposed and developed a theoretical model for describing single scattering based on ellipsoidal coordinate systems and evaluated various combinations of absorption and scattering coefficients, anisotropic phase functions, and transmitter/receiver distances [3].

In 1985, Geller et al. [4] used a mercury lamp as the ultraviolet light source and a photomultiplier tube and filter combination as the receiver to verify the feasibility of ultraviolet communication. In 1991, Luettgen et al. [5] proposed a single-scattering model describing uniform scattering and dielectric absorption. They analyzed the angular spectrum of single-scattered energy, as well as the pulse response time and path loss of short-range indirect communication systems.

In 2000, the General Telephone and Electronics (GTE) cooperated with the US military to develop ultraviolet optical communication systems. They achieved an

indirect communication distance of less than 2 km, an information transmission rate of 4.8 kbit/s, and a bit error rate of 10^{-6}, which met the requirements of video transmissions [6].

Early ultraviolet communication systems mostly used gas-discharge lamps as light sources, which had the disadvantages of a nonunique beam wavelength, small modulation bandwidth, and large size. To miniaturize this system, the Defense Advanced Research Projects Agency (DARPA) launched the Semiconductor UV Optical Source (SUVOS) program in 2002 and the Deep UV Avalanche Photodiode Program in 2007 [7].

In 2005, Shaw et al. [8] developed a modular solar-blind ultraviolet communication experiment bench and conducted field experiments in various environments, such as open areas and under a canopy. The measured data and supported single-scatter models showed that solar-blind UV light-emitting diodes (LEDs) could play an important role in short-range communication links with NLOS and LOS.

In 2008, Xu et al. [9] used commercial UV LEDs as a light source to build an ultraviolet communication test bench to evaluate the path loss and bit error rate of a communication system. When using on-off keying (OOK) modulation, the bit error rate of the communication system was obtained by changing the angles of the receiver (Rx) and transmitter (Tx), with different Rx elevation angles. When the Rx elevation angle was 30°, the bit error rate increased from 10^{-6} to 10^{-1}.

In 2009, Majumdar et al. [10] used the Monte Carlo simulation of photon tracking to establish a random indirect UV communication (UVC) channel model and compared multiple and single scattering. Indirect UV-scattering channels were studied from the aspects of impulse response, path loss, bandwidth, etc.

In 2011, Xu et al. [11] compared the path loss and bit error rate (BER) performance of links in different environments based on original commercial UV LEDs. By considering the unique UV physical layer characteristics, that is, changing the channel quality along different scattering directions, an efficient neighbor discovery protocol was developed. In 2013, the team proposed a configurable framework that combined the traditional line-of-sight turbulence theory with a Monte Carlo simulation framework for random photon scattering. The results showed that the model effectively characterized the irradiance probability density function (PDF) of the received signal of a UV channel under strong and weak turbulence [12].

In 2015, Mohamed et al. [13] proposed a spatial diversity technique for delay compensation and detector-noise suppression in indirect UV optical communication. The advantages of interunit receivers were compared and analyzed from the perspective of the information transmission rate. Experiments showed that, under the same signaling conditions, the use of diversity technology increased the information transmission rate of the single wire receiver by 50%.

In 2016, Ardakani et al. [14] conducted a simulation study on an NLOS multi-input–multi-output (MIMO) communication system and analyzed the relationship between the signal-to-noise ratio and bit error rate in three different modes: NLOS ultraviolet MIMO, multiple-input single-output, and single-input multiple-output.

In 2017, Maryam et al. [15] proposed a multihop UV system with decoding and forwarding relays, considering the turbulent influence of short-distance communication. The performance of the working link under lognormal turbulence was analyzed, including the probability density function of the interrupt probability, diversity gain, and power attenuation factor at different atmospheric refractive index structure constants and communication distances.

In 2018, Refaai et al. [16] analyzed the performance of multiple NLOS UV communication cooperative relays in turbulent channels. They used decoding and forwarding relay technology to complete the NLOS UV communication according to the channel state information (CSI). The interruption probability of a multirelay UV system was discussed according to the proposed model, and the system performance under different elevation angles, atmospheric-scattering parameters, and receiver field-of-view angles was simulated and analyzed. The experimental results showed that increasing the number of UV indirect coordinated relays did not necessarily improve the system performance; other factors, such as the elevation angle and the number of relays, must also be considered.

In 2019, Arya et al. [17] proposed a multihop ultraviolet NLOS communication system based on amplification and forwarding and adopted subcarrier intensity modulation to mitigate the impact of turbulence. Subsequent simulation results showed that the system could effectively mitigate turbulence-induced fading without sacrificing important performance criteria, such as the interrupt probability and average bit error rate.

In 2020, Hasan et al. [18] derived a system model under multiple-input single-output, single-input multiple-output, and MIMO modes by studying the spatial diversity reception technology in NLOS ultraviolet optical communication. They analyzed the bit error rate performance of the communication link in the single-input single-output system model and verified the accuracy of the model with Monte Carlo simulation.

In 2021, Bhatia et al. [19] characterized the expression for the interruption probability of a nonlinear UV communication system by analyzing the multiplicative distortion statistics caused by turbulence. Using the proposed multiplicative distortion model, the bit error rate of the nonlinear outdoor indirect UV channel was quantified, and an equalizer based on the minimum symbol error rate (MSER) criterion of the reproducing kernel Hilbert space was proposed, according to the derived bit error rate expression, to reduce the distortion caused by LED nonlinearity.

In 2022, Arslan et al. [20] proposed a novel steering optimization method based on a finite difference stochastic approximation (FDSA) to simultaneously optimize the pointing direction of the transmitter (Tx) and receiver (Rx) without knowing the position or relative orientation of the two nodes. Monte Carlo channel simulation was used for the parameter analysis to study and select the appropriate key algorithm parameters. Experiments with custom-designed UV Tx and Rx gimbal systems demonstrated the practicality and efficiency of the proposed steering optimization method, and the received photon count was significantly increased.

Domestic Research Status

Major domestic research institutions include the Beijing Institute of Technology, Xi'an University of Technology, Chinese Academy of Sciences, and other institutions. Most studies are still in the prototyping and feasibility testing stages, and some research on channel models is still being explored.

In 1999, Ni et al. [21] used a low-pressure gas ultraviolet mercury lamp with a 253-nm wavelength as the light source and a photomultiplier tube to build an ultraviolet communication experimental platform for the receiver. They completed voice communication at a distance of 500 m. In 2003, the team used a low-transmittance model to simulate the atmospheric transmission characteristics of the ultraviolet band and analyzed the influence of the elevation difference, visibility, weather, transmission distance, and other factors on the atmospheric ultraviolet transmittance at the transceiver end. The results showed that around the 250-nm band, weather, visibility, season, and other factors had little influence on the ultraviolet transmittance of the solar-blind zone [22].

In 2004, the half-duplex ultraviolet communication system built by Changsheng et al. [23] realized 9.6-kbit/s voice communication, laying the foundation for subsequent ultraviolet optical communication.

In 2005, Lan et al. [24] proposed the use of frequency modulation to realize language coding in ultraviolet light information transmission based on the ultraviolet communication system of the General Telephone and Electronics (GTE). In the experiment, an ultraviolet germicidal lamp was used as the light source, and a solar-blind photomultiplier tube was used as the photodetector. It realized a data transmission rate of 1.2 kbit/s, laying a good foundation for solar-blind ultraviolet communication systems.

In 2007, Xiao [26] used a low-voltage mercury lamp as the light source for an ultraviolet optical communication system and analyzed the shortcomings of the device, such as its small modulation bandwidth and high power consumption. Subsequently, ultraviolet LEDs replaced the low-voltage mercury lamps, and pulse-position modulation (PPM) technology was used to create an experimental prototype with a communication distance of 10 m [25]. In 2012, a field-programmable gate array (FPGA) was used for modulation and demodulation, realizing 9.6-kbit/s coded communication and 115-kbit/s noncoded communication.

In 2010, Ke [27] proposed to use the multihop feature of self-organizing networks to overcome the shortcomings of the limited ultraviolet transmission distance. He added a reward and punishment mechanism to improve the original binary exponential retreat algorithm. The competitiveness of the nodes matched their competitive windows, thereby achieving fair access to channels.

In 2011, based on the characteristics of NLOS ultraviolet communication, Ke et al. derived the coverage azimuth expression, improved the NLOS directional transmission node structure, and proposed and simulated an NLOS directional medium access control (MAC) protocol based on angle perception [28]. The results showed that the protocol greatly improved the network throughput and reduced the

average communication delay. In 2012, the Monte Carlo (MC) multiscattering ultraviolet optical channel model was further studied; based on the single-scattered channel model, the node coverage of short-range communication was analyzed, and its mathematical expression was derived [29].

In 2013, Wu et al. [30] designed and implemented a MAC layer power control mechanism suitable for ultraviolet optical communication. It was designed to reduce the energy consumption of nodes and mutual interference between nodes to extend the life cycle of the network and increase the system capacity. In the same year, Han et al. [31] adopted a low-density parity check code algorithm to improve the system performance. They improved the communication distance without coding by 78% under a rated transmit power and a bit error rate of 10^{-5}.

In 2015, Yang et al. [32] proposed a UV directional neighbor discovery algorithm, based on traditional time-division multiple access, that was applied to high-speed large-scale mobile UV ad hoc networks. Compared with the traditional algorithm, the probability of neighbor node discovery was increased by 45%, and node collisions were reduced; however, the convergence speed of the algorithm was sacrificed.

In 2015, Li et al. [33] proposed an anti-jamming relay link scheme and analyzed the corresponding power requirements considering that the ultraviolet optical communication distance was too short, owing to limited emission optical power and serious atmospheric attenuation. The scheme allocated different time slots to adjacent nodes through time-division multiplexing to prevent mutual interference between nodes.

In 2016, Song et al. [34] simulated and analyzed the interference between links in space-division multiplexing, based on a set of typical communication link models using the MC method. In addition, the pulse-broadening problem in the noncoplanar case was further analyzed. In 2018, a mobile self-organizing communication node device integrating transmission and reception was designed, and the relationship between capture time and speed was simulated and analyzed [35].

In 2017, Li et al. [36] established a high-altitude solar radiation distribution model to analyze high-altitude ultraviolet optical communication performance. They studied the high-altitude scattering and absorption coefficients of ultraviolet light of different wavelengths and simulated the loss and signal-to-noise ratio of direct and indirect UV links, considering the background light of solar radiation and shot particle noise at the receiving end.

In 2019, Wang et al. [37] proposed the use of multidecimal frequency-shift keying for modulation, in view of the low modulation rate, and developed an ultraviolet optical communication system for experiments. Using a quaternary frequency-shift keying modulation method, the maximum transmission rate reached 30 kbit/s, which was double that of the binary frequency-shift keying method.

In 2020, Zhang et al. [38] built an FPGA-based ultraviolet optical communication system and analyzed its time delay. The results showed that using an FPGA to build a communication system effectively reduced the communication delay.

In 2021, Cao Tong et al. [39] proposed a single-scattering loss model for an NLOS UVC system based on ultraviolet LEDs. They assumed that the LED

emission model followed a Lambertian distribution and derived the corresponding analytical expression. Monte Carlo simulations and experimental results verified the effectiveness of the proposed loss model. The experimental results demonstrated that the proposed model accurately predicted the loss of single scattering under actual conditions. Compared to the actual measurement results, the root-mean-square error of the proposed loss model was approximately 1 dB.

In 2021, Cheng et al. [40] proposed a UV lossless competition MAC protocol, considering the shortcomings of the UV MAC protocol. Based on the superposition logic of UV power, the MAC protocol effectively reduced the collision loss of multinode competing access. The basic working mechanism and protocol flow were provided in detail, and the network performance was mathematically simulated and analyzed, based on a four-node UV network. Simulation results showed that, compared with the time-slot ALOHA protocol, the protocol had a higher throughput and lower latency. Subsequently, the actual networking of the four communication terminals was designed and realized, and the effectiveness of the new UV-MAC protocol was fully verified.

In 2022, Cao et al. [41] established a reflection-assisted NLOS ultraviolet communication path loss analysis model combining a single collision-induced event (CIE) by treating both scattering and reflection events as CIEs, which referred to a single collision-induced (SCI) path loss model. Monte Carlo (MC) simulations and experimental results verified the correctness and validity of the proposed SCI path loss model. The results showed that the proposed SCI path loss model worked well for both coplanar and noncoplanar geometries of reflection-assisted NLOS UVC.

Theoretical Model of Ultraviolet Background Light

Because of the absorption of the ozone layer, the intensity of ultraviolet radiation in the 200-nm–280-nm band near the surface is very low; this is called the solar-blind zone [42]. Most current research focuses on the central band of the surface solar-blind zone, with low background noise. To reduce the influence of background noise on communication quality and achieve more stable high-altitude ultraviolet communication and a wider communication bandwidth, it is necessary to study the background noise of ultraviolet light in the solar-blind zone.

Theoretical Model of Solar Radiation to the Earth

Total solar radiation calculation models can be divided into three categories: physical, empirical, and machine learning [43]. Among them, the empirical model is one of the most widely used models because of its high computational accuracy and low computational complexity. Angstrom [44] first established a calculation model for total daily solar radiation, based on the correlation between the insolation percentage and total radiation/sunny day radiation. However, because of the complexity of the process of calculating sunny day radiation, Prescott [45] proposed the use of

astronomical radiation instead of sunny day radiation and established a model based on the Angstrom model, also known as the A-P model.

Photon Energy

The Sun transmits energy in the form of light throughout the solar system and beyond. The Sun's photons carry the Sun's energy; they are light quanta, the quanta of electromagnetic radiation, and the normative particles that transmit electromagnetic interactions. A photon's resting mass is zero, it has no charge, and its energy is the product of Planck's constant and the frequency of electromagnetic radiation, which can be expressed by the following equation [46]:

$$E = \frac{hc}{\lambda} = hf, \tag{45.1}$$

where h is Planck's constant, $h = 6.6260693 \times 10^{-34}$J. s, c is the speed of light, $c = 2.998 \times 10^{8}$m/s, and f is the frequency of light at different wavelengths. Light composed of high-energy photons has a shorter wavelength, while light composed of low-energy photons has a longer wavelength.

The amount of photon energy can be expressed as electron volts, in addition to joules. The electron volt represents the energy required to raise an electron by one volt; thus, a photon has the energy of $1\text{eV} = 1.602 \times 10^{-19}$J. The above constant hc can therefore be recharacterized in eV [47].

$$hc = \left(1.99 \times 10^{-25}\text{J.m}\right) \times \left(1\text{eV}/1.602 \times 10^{-19}\text{J}\right) = 1.24 \times 10^{-6}\text{eV.m} \tag{45.2}$$

The photon energy equation is transformed into eV and μm, and a common expression is obtained, which represents the relationship between photon energy and the wavelength, as shown below:

$$E(\text{eV}) = \frac{1.24}{\lambda(\mu\text{m})}. \tag{45.3}$$

The photon energy of different bands is different. Table 45.1 shows the frequency, wavelength, and photon energy of photons of different bands.

As can be seen from the table, as the wavelength decreases, the energy carried by the photon gradually increases, that is, the shorter the wavelength, the greater the

Table 45.1 Photon energies for different wavebands [48]

Band	Frequency f/Hz	Wavelength λ/μm	Photon energy hf/eV
Microwaves	3×10^{8}–3×10^{12}	10^{6}–10^{2}	1.2×10^{-6}–1.2×10^{-2}
Infrared light	8.8×10^{11}–4.3×10^{14}	3.4×10^{2}–7×10^{-1}	3.6×10^{-3}–1.7
Visible light	4×10^{14}–7.5×10^{14}	7.6×10^{-1}–4×10^{-1}	1.6–3
Ultraviolet light	7.5×10^{14}–3×10^{16}	4×10^{-1}–10^{-2}	3–120
X-rays	3×10^{16}–3×10^{20}	10^{-2}–10^{-6}	1.2×10^{2}–1.2×10^{6}
Gamma (γ) rays	$>3 \times 10^{19}$	$>10^{-5}$	$>1.2 \times 10^{5}$

energy of the photon. The minimum energy of a microwave is 1.2×10^{-6} eV–1.2×10^{-2} eV, and the maximum energy of a γ ray is more than 1.2×10^{5} eV.

Blackbody Radiation Model

Blackbody radiation is a thermal radiation phenomenon, that is, the law of change in the intensity and frequency of the radiation spectrum emitted by an object at different temperatures. The origin of this theory can be traced back to the early nineteenth century. The law of blackbody radiation was proposed by Planck and Wien and provided the basis for a large number of subsequent physics and astronomical studies [49].

Suppose that a black body is composed of many simple harmonic oscillators of different vibrational frequencies and energy level and that each harmonic oscillator can emit or absorb radiation of a certain frequency (energy). For a harmonic oscillator in an energy state whose energy is E, according to the Boltzmann distribution function, the number of simple harmonic oscillators for that energy state is [50]:

$$n(E) = \frac{g(E)}{\exp\left(\frac{E}{\kappa_B T}\right) - 1}, \tag{45.4}$$

where $g(E)$ is the number of harmonic oscillators in the blackbody at each energy level, κ_B is the Boltzmann constant, $\kappa_s = 1.380 \times 10^{-23}$J/K, and T is the blackbody temperature. At very high temperatures, the -1 term in the denominator can be ignored:

$$n(E) \approx \frac{g(E)}{\exp\left(\frac{E}{\kappa_B T}\right)}. \tag{45.5}$$

Within a certain frequency range, all harmonic oscillators located between energy E and $E + \Delta E$ can emit all the electromagnetic waves with energy of ΔE. Therefore, all electromagnetic waves at that frequency should also have the energy ΔE. Let this frequency range be v to $v + \triangle v$, so the energy flow through this band in a unit time (i.e., the radiation flux density) can be expressed as

$$I_v(v, T) = \frac{g(E)}{\exp\left(\frac{E}{\kappa_B T}\right) - 1} \cdot \frac{hv^2}{\lambda^3}. \tag{45.6}$$

According to electromagnetic wave theory, converting frequency v to wavelength λ yields

$$I_\lambda(\lambda, T) = \frac{hc^2}{\lambda^5 . \exp\left(\frac{hc}{\lambda \kappa_B T}\right) - 1}, \tag{45.7}$$

where λ is the wavelength of light, T is the blackbody temperature in K, h is Planck's constant, $h = 6.626 \times 10^{-34}$J. s, and c is the vacuum speed of light, 3×10^8 m/s.

This is a basic formula used in the fields of cosmology, astrophysics, and infrared technology.

Solar Radiation Model

Figure 45.1 shows a schematic representation of solar radiation from the Sun to the Earth's surface, where β is the ecliptic angle, the direction of the vernal equinox is 0° longitude, and α is called the yellow meridian of the Sun and is the angle between the current line of the Earth and the Sun to the line of the Earth and the vernal equinox. ν_0 is the meridian of perihelion. The axes (X, Y, Z) represent the coordinate system of the equator. It can be transformed by the zodiacal coordinate system (X', Y', Z') according to Eq. (45.8) [51]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \tag{45.8}$$

The position of the Sun can be expressed as P_s^y, a function of the day's longitude α: $P_s^y = [r\cos(\alpha) \;\; r\sin(\alpha) \;\; 0]^T$. r is the distance from the Sun to the Earth, which can be expressed as

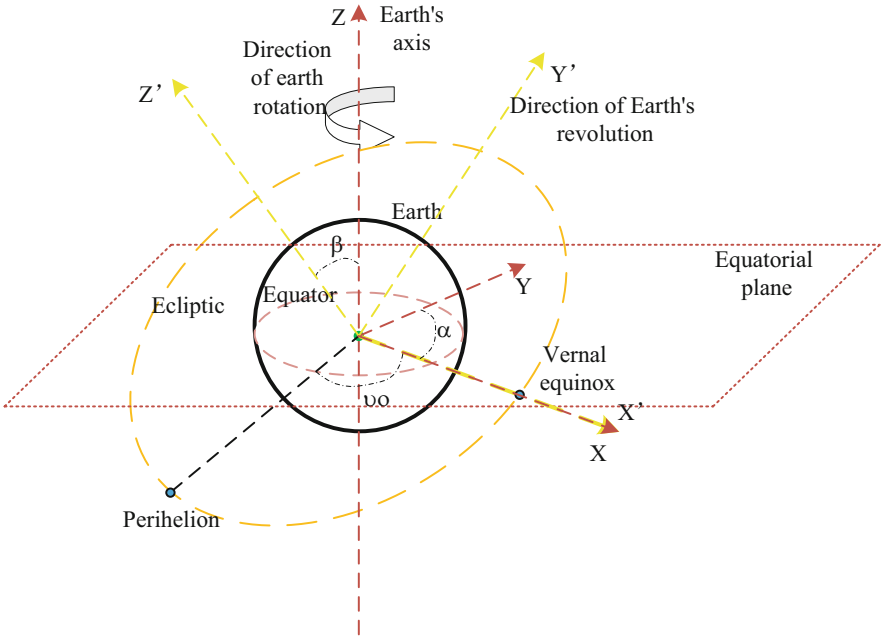


Fig. 45.1 Schematic diagram of solar radiation to the Earth

$$r = A\left(1 - \frac{e(\cos(\alpha + v_0))}{1 + e\cos(\alpha + v_0)}\right), \tag{45.9}$$

where A is the semimajor axis of the Earth's orbit, $A = 1.496 \times 10^{11}$ m, and e is the eccentricity of the solar orbit, $e = 0.016722$. The ecliptic α for a given date can be derived by the following formula:

$$\frac{\int_0^{\alpha} \frac{1}{2} r^2 d\alpha}{\pi A\sqrt{1 - e^2}} = \frac{n}{365}, \tag{45.10}$$

where n is the cumulative day, $\int_0^{\alpha} \frac{1}{2} r^2 d\alpha$, the number of days away from the vernal equinox, $\pi A\sqrt{1 - e^2}$ is the Earth–Sun line from the vernal equinox to the current date across the area, and A is the area of the ecliptic ellipse.

When there is an angle γ between the solar radiation and the surface horizontal normal vector, as shown in Fig. 45.2, the total surface solar radiation intensity can be expressed as [52]

$$E_0 = \frac{M_0}{4\pi r^2} \cos\gamma, \tag{45.11}$$

where M_0 is the external solar radiant power with a constant of 3.9×10^{26} W. $\cos\gamma$ can be given by the following equation:

$$\cos\gamma = \frac{P_s^r n_g^r}{\left\|P_s^r\right\|}, \tag{45.12}$$

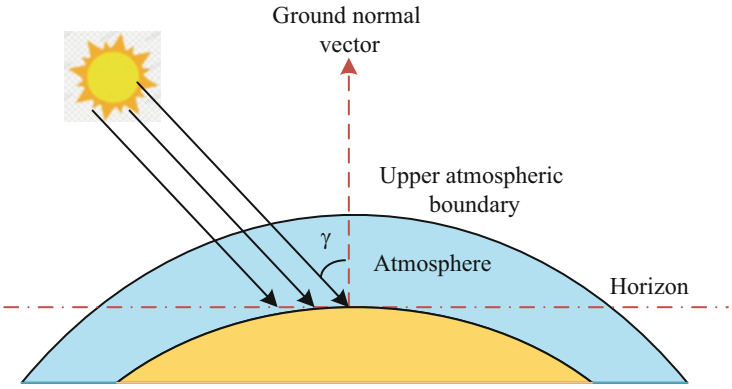


Fig. 45.2 Solar radiation on the ground

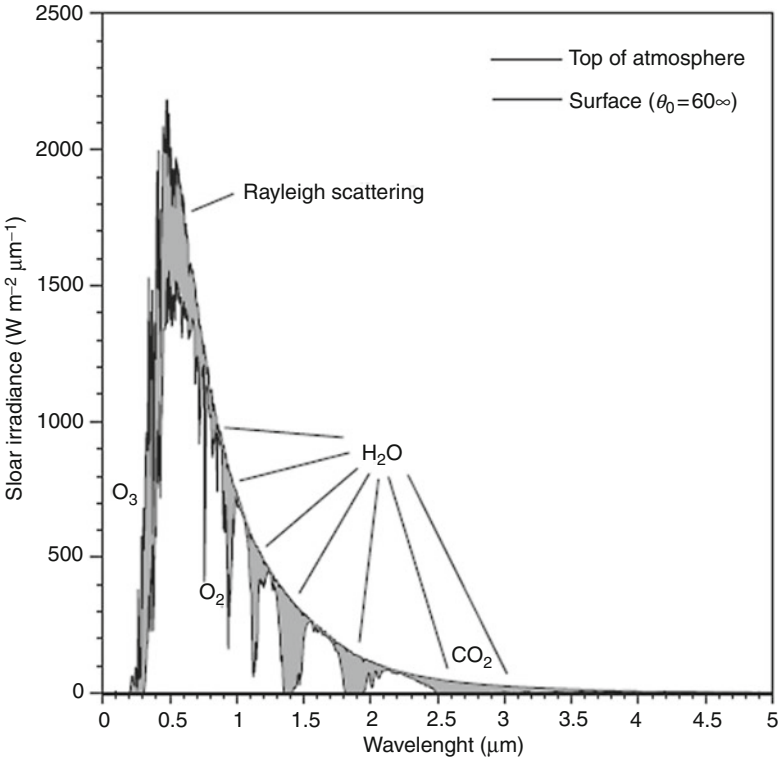


Fig. 45.3 Solar spectrum on a clear day [55]

where P_s^r is the position of the Sun in the equatorial system, which can be derived from the transformation matrix in (45.8); n_g^r is the surface normal vector expressed as $n_g^r = [\cos\theta\cos\phi \quad \sin\theta\sin\phi \quad \sin\phi]^T$, ϕ is the local latitude, and θ is the right ascension. Then, the solar radiation intensity can be expressed as [53, 54]:

$$E_0 = \frac{M_0}{4\pi r^2 \left\| P_s^r \right\|} P_s^r n_g^r. \tag{45.13}$$

Figure 45.3 shows the solar spectrum of the Earth's surface at a zenith angle of 60° in clear weather conditions [55].

Atmospheric Parameters

Atmospheric transparency is an important feature of the radiation climate and a decisive factor for direct ultraviolet radiation on the surface. It not only objectively reflects the atmospheric transparency and atmospheric physical state of optical radiation but is also an important parameter for characterizing the attenuation degree of ultraviolet radiation by the atmosphere [56].

For a certain solar altitude angle, changes in atmospheric transparency and clouds are the main factors that determine the intensity of ultraviolet radiation [57]. Direct radiation is weakened by the reflection and absorption of clouds. However, clouds influence ground long-wave effective radiation through two main aspects: First, the shielding effect of clouds on short-wave direct radiation reduces surface heating, thereby weakening long-wave ground radiation. Second, the presence of clouds increases atmospheric inverse radiation, such that the effective radiation on the ground is weakened to a greater extent [58].

According to the Bouguer–Lambert theorem, when solar radiation I with a wavelength A passes through an atmosphere dm thick, its radiative decay is reduced to [59]:

$$dI_\lambda = -\alpha_\lambda I_0.\lambda dm, \tag{45.14}$$

where α_λ is the extinction coefficient of the atmosphere.

For the solar-blind zone band in daylight, there should be a solution for the entire band from 200 nm to 280 nm:

$$S_m = \int_{0.2}^{0.28} S_0, \lambda P_\lambda^m d\lambda. \tag{45.15}$$

It can be simplified to

$$S_m = S_0 P_\lambda^m \tag{45.16}$$

$$P_\lambda = \sqrt[m]{S/S_0}, \tag{45.17}$$

where P_λ is the compound transparency coefficient, m is the atmospheric mass, and S_0 is the solar constant, $S_0 = 1381.6W.\, m^{-2}$. S_m indicates the solar radiation flux on the ground perpendicular to the surface of the Sun's rays when the atmospheric mass is m.

Equation (45.17) is an exact theoretical equation, not an empirical relationship. P_λ^m increases with m, a phenomenon known as the Forbes effect. The application of this effect greatly limits P_λ. To overcome the Forbes effect, atmospheric transparency must be adjusted to a given atmospheric quality. That is, $m = 2$ describes the calculated integral transparency. P can be divided into six grades: very turbid ($P < 0.625$), turbid ($P = 0.625$–0.675), low transparent ($P = 0.676$–0.725), normal transparent ($P = 0.726$–0.775), high transparent ($P = 0.776$–0.825), and very transparent ($P \geq 0.826$) [60].

According to the measurement data of meteorological stations in various places, China has established P_λ, an empirical check table that is corrected to P. The revised formula derived from the check table is

$$P = (1.91/S_0)^{1/2}(S_h/1.91)^{(\sin h+0.15)/1.3}, \tag{45.18}$$

where h is the solar altitude angle and S_h is the direct solar radiation perpendicular to the solar ray plane.

Solar-Blind Zone UV Background Light Radiation Model

The energy of solar radiation is primarily concentrated in the wavelength range of 0.15 μm–4 μm. This wavelength range can be divided into three main regions: the ultraviolet region with shorter wavelengths, the infrared region with longer wavelengths, and the visible region in between. Solar radiation is mainly distributed in the visible and infrared regions, with the former accounting for 50% of the total solar radiation and the latter accounting for 43%. The ultraviolet region accounts for only 7% of the energy. The capacity of solar radiation reaches its highest value at a wavelength of 0.475 microns.

The radiant capacity of the various wavelengths of the sun can be described by the blackbody radiation model, in which $w(\lambda)$ can be expressed as [61]

$$w(\lambda) = \frac{c^2 h}{\lambda^5}\left\{\frac{1}{\exp[hc/(k\lambda T)] - 1}\right\}, \tag{45.19}$$

where c is the vacuum speed of light, $c = 3 \times 10^8$ m/s; h is Planck's constant, $h = 6.626 \times 10^{-34}$ J s; T is the temperature in Kelvin; and k is Boltzmann's constant, $k = 1.380 \times 10^{-23}$J/K. Then, the percentage of the overall radiation capacity of a certain band of ultraviolet $F_{(\lambda_1-\lambda_2)}$ can be expressed as [62]

$$F_{(\lambda_1-\lambda_2)} = \frac{\int_{\lambda_2}^{\lambda_1} w(\lambda)d\lambda}{\int_0^{\infty} w(\lambda)d\lambda} = \frac{1}{\sigma T^4}\int_{\lambda_2}^{\lambda_1} w(\lambda)d\lambda = \frac{1}{\sigma T^4}\left[\int_0^{\lambda_1} w(\lambda)d\lambda - \int_0^{\lambda_2} w(\lambda)d\lambda\right]. \tag{45.20}$$

For solar-blind ultraviolet communication systems, the sensitive wavelength range of background noise is 200 nm–300 nm, and the main source is sunlight [63]. Figure 45.4 shows the blackbody radiation pattern when the surface temperature of the Sun is 5800 K. The gray area is the ultraviolet background light radiation spectrum, and its integral sum is the sum of the solar background noise in this band.

At latitude ϕ, longitude θ, receiver radius $D_r/2$, filter transmittance η_B, and receiving field of view $\theta_{r,\text{signal}}$, the receiving end of the UV background light power is [64]

$$P_B = F_{(\lambda_1-\lambda_2)} \cdot \eta_B \pi \left(\frac{D_r}{2}\right)^2 E_0 \cdot \left[P^m + \frac{1 - P^m}{2(1 - 1.4\ln P)}\right] \cdot \frac{\pi}{4}\theta_{r,\text{signal}}^2, \tag{45.21}$$

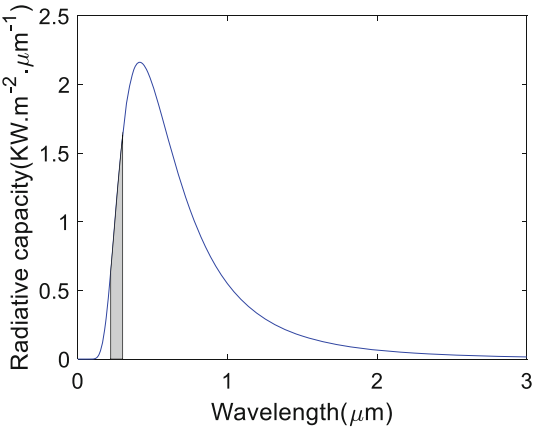


Fig. 45.4 Solar radiation capacity at each wavelength

where $F_{(\lambda_1-\lambda_2)}$ represents the percentage of a certain band to the overall radiation capacity, η_B is the filter transmittance, D_r is the receiver diameter, E_0 is the surface solar radiation intensity, P is the atmospheric transparency, and m is the air quality.

Assume that the measurement time is from 6 a.m. to 6 p.m. on a sunny day and the measurement location is Xi'an, China. The measurement wavelength is 300 nm, the receiver diameter is 10 mm, the receiving field of view is 20 mrad, the filter transmittance is 0.24, and the bandwidth is $\pm$10 nm. The winter solstice in Xi'an was simulated on December 22, 2022; the spring equinox on March 21, 2023; the summer solstice on June 21; and the autumn equinox on September 23.

A schematic diagram of the four solar terms, variation of ultraviolet background light, and days under sunny conditions with the seasons are shown in Fig. 45.5. It can be seen that the background light radiation value of the summer solstice is the largest, those of the autumn and spring equinoxes are similar, and the winter solstice is the smallest. With the change in seasons, the peak daily radiation time will also change. Figure 45.6 shows a comparison of the measured and theoretical values of the winter solstice, spring equinox, summer solstice, and autumn equinox. It can be seen that the winter solstice, spring equinox, summer solstice, and autumn equinox correspond to the theoretical changes in Japanese and Chinese times.

UV Optical Communication Path Loss Model

Single-Scattering Model and Theory

The single-scattering model, proposed by Luettgen et al. [65] in 1991, is an extension of the single-scattering transceiver axis collinear model proposed by Reilly and Warde [66] in 1979, which is essentially a single-scattering coplanar direct integration model, that is, the transceiver axis is coplanar. To facilitate the calculation, the ellipsoid coordinate system shown in Fig. 45.7 is established, and each point in the figure can be uniquely determined by the diameter ξ, corresponding angle η of the radial quantity, and azimuth φ.

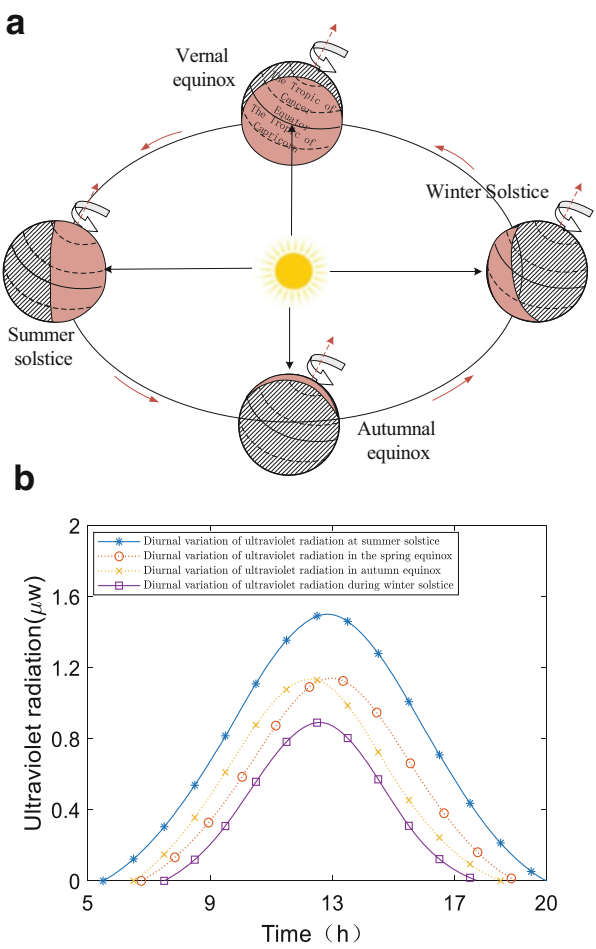


Fig. 45.5 Diurnal variation of four solar terms and ultraviolet background radiation. (**a**) Schematic of the four solar terms. (**b**) Theoretical values of the diurnal variation in ultraviolet radiation in different seasons

As shown in Fig. 45.7, F_1 and F_2 are the two focal points of the ellipsoid, r is the focal length of the ellipsoid, r_1 and r_2 are the two focal radii of the ellipsoid, and ψ_1 ψ_2 are the two focal angles of the ellipsoid.

The transformation relationship of any point under the Cartesian coordinate system (x, y, z) to the ellipsoid coordinate system (ξ, η, φ) is

$$\xi = (r_1 + r_2)/r \ \ (1 \le \xi \le \infty) \tag{45.22}$$

$$\eta = (r_1 - r_2)/r \ \ (\text{-}1 \le \eta \le 1) \tag{45.23}$$

$$\phi = \arctan(y, x) \ \ (\text{-}\pi \le \xi \le \pi) \tag{45.24}$$

$$\theta_s = \psi_1 + \psi_2. \tag{45.25}$$

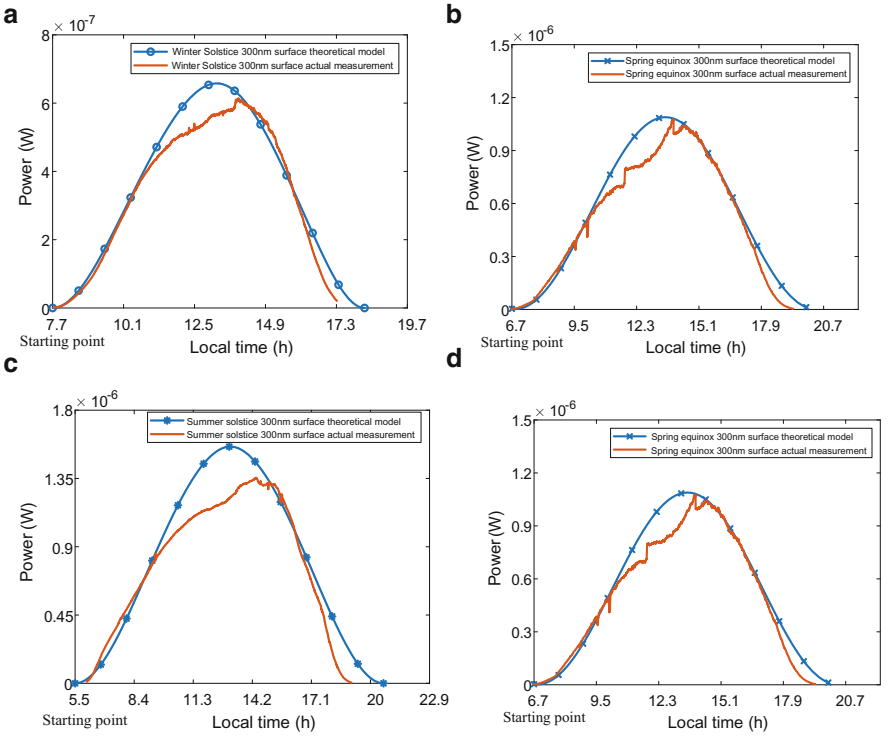


Fig. 45.6 Measured and theoretical values of the winter solstice, spring equinox, summer solstice, and autumn equinox. (**a**) Winter solstice. (**b**) Spring equinox. (**c**) Summer solstice. (**d**) Autumn equinox

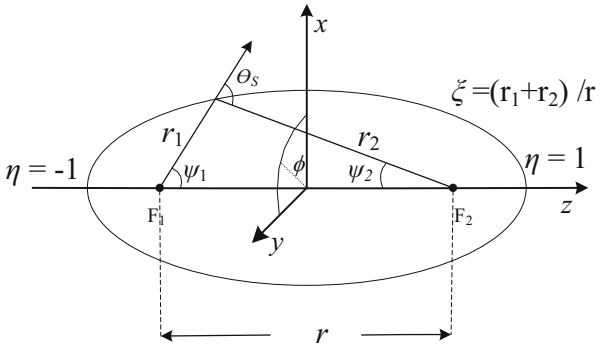


Fig. 45.7 Single-scattering ellipsoidal coordinate system

Equations 45.22 and 45.23 can be solved for r_1 and r_2:

$$r_1 = \frac{r(\xi + \eta)}{2} \tag{45.26}$$

$$r_2 = \frac{r(\xi - \eta)}{2}. \tag{45.27}$$

In an ellipsoidal coordinate system, the focal and scattering angles can be expressed as

$$\cos\psi_1 = (1 + \xi\eta)/(\xi + \eta) \tag{45.28}$$

$$\sin\psi_1 = \sqrt{(\xi^2 - 1)(1 - \eta^2)}/(\xi + \eta) \tag{45.29}$$

$$\cos\theta_s = \left(2 - \xi^2 - \eta^2\right)/\left(\xi^2 - \eta^2\right). \tag{45.30}$$

The single-scattering model is simplified by assuming that photons are emitted from the transmitting end and scattered only once before their arrival at the receiving end. The single-scattering coplanar model is shown in Fig. 45.8. The receiver and transmitter are located on the two focal points of the ellipsoid, F_1 and F_2, respectively; θ_R represents the receiving field-of-view half-angle; θ_T is the transmission field-of-view half-angle; β_R is the receiving elevation angle; β_T is the emission elevation angle; and θ_s is the scattering angle; $\theta_s = \beta_T + \beta_R$.$r$ is the distance between the transmitter and the receiver; r_1 is the distance from the transceiver to the effective scattering point; and r_2 is the distance from the effective scattering point to the receiver. The shaded part is the effective scattering volume V, that is, after one scattering, only the atmospheric light signal located in volume V can be received by the receiver.

It was assumed that the emission and receiving optical axes were in the same plane, and the influence of atmospheric turbulence was not considered. At moment

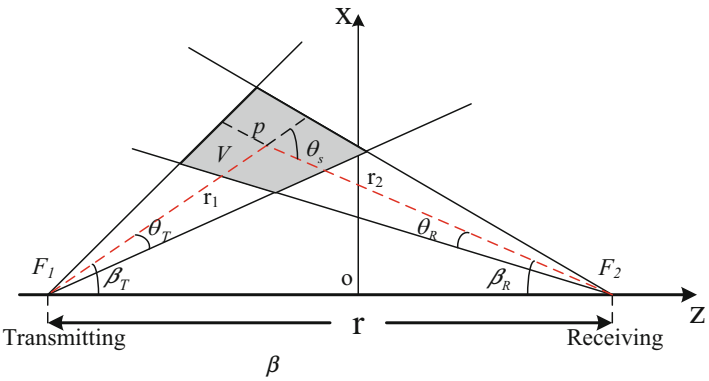


Fig. 45.8 Single-scattering NLOS model

$t = 0$, the emitter emits a pulse of ultraviolet light with an impulse response energy E_T and leaves the emitter at emission elevation angle β_T. After scattering and absorption by the atmosphere, it reaches point $p = (\xi, \eta, \varphi)$ at a distance r_1 from the emitter. $t = r_1/c$, where c is the vacuum speed of light. The energy at point p is

$$H_p = \frac{E_T \exp(-k_e r_1)}{\Omega_r \left(r_1^2\right)}, \tag{45.31}$$

where $\Omega_r = 2\pi(1 - \cos\theta_T)$ is the inherent cone angle of emission and k_e is the atmospheric extinction coefficient.

The differential volume element dV can be thought of as a secondary light source and can be seen as a spherical point light source that emits the following energy:

$$dR_p = k_s \frac{E_T \exp(-k_e r_1)}{\Omega_r \left(r_1^2\right)} \cdot \frac{P(\cos\theta_s)}{4\pi} dV. \tag{45.32}$$

The energy dR_p reaches the receiver after transmission distance r_2, and the energy received per unit area of the receiver is

$$\begin{aligned} dH_R &= dR_p \cdot \frac{\cos(\xi) \exp(-k_e r_2)}{r_2^2} \\ &= \frac{E_T k_s \cos(\xi) \exp(-k_e (r_1 + r_2))}{4\pi \Omega_r (r_1 r_2)} P(\cos\theta_s) dV, \end{aligned} \tag{45.33}$$

where k_s is the angle between the line of the scattering point s and ξ, the center of the receiving surface of the receiving end. $P(\cos\theta_s)$ is a scattering phase function, and $\cos(\xi)$ is the center axis of the cone of the receiving field of view:

$$\cos(\xi) = \sin\beta_R \sin\psi_1 \cos\phi + \cos\beta_R \cos\psi_1. \tag{45.34}$$

In ellipsoidal coordinates, dV can be expressed as

$$dV = \frac{r^3}{8} \left(\xi^2 - \eta^2\right) d\xi d\eta d\phi. \tag{45.35}$$

Substituting Eqs. (45.26), (45.27), and (45.35) into Eq. (45.32) yields

$$dH_R = \frac{E_T k_s \cos(\xi) \exp(-k_e r \xi)}{2\pi \Omega_r r \left(\xi^2 - \eta^2\right)} P(\cos\theta_s) d\xi d\eta d\phi. \tag{45.36}$$

The time for the ultraviolet light pulse to reach the receiver $t = (r_1 + r_2)/c$ and $\xi = (r_1 + r_2)/r$.

$$dE = \frac{E_T c k_s \cos(\xi) \exp(-k_e r \xi)}{2\pi \Omega_r r^2 \left(\xi^2 - \eta^2\right)} P(\cos\theta_s) d\eta d\phi \tag{45.37}$$

At moment $t = \xi r/c$, the integral of Eq. (45.36) provides an acceptable irradiance in W/m^2, which can be expressed as

$$E = \begin{cases} \dfrac{E_T c k_s \exp(-k_e r\xi)}{2\pi\Omega_r r^2} \displaystyle\int_{\eta_1(\xi)}^{\eta_2(\xi)} \int_{\phi_1(\xi,\eta)}^{\phi_2(\xi,\eta)} \dfrac{\cos(\xi) P(\cos\theta_s)}{\xi^2 - \eta^2} d\eta d\phi & (\xi_{\min} < \xi < \xi_{\max}) \\ 0 & (\text{otherwise}) \end{cases} \tag{45.38}$$

From the symmetry of ellipsoidal coordinates, $\phi_2(\xi, \eta) = -\phi_1(\xi, \eta)$; integrating with regard to φ, Eq. (45.37) can be reduced to

$$E = \begin{cases} \dfrac{E_T c k_s \exp(-k_e r\xi)}{2\pi\Omega_r r^2} \displaystyle\int_{\eta_1(\xi)}^{\eta_2(\xi)} \dfrac{2g[\phi_2(\xi,\eta)] P(\cos\theta_s)}{\xi^2 - \eta^2} d\eta & (\xi_{\min} < \xi < \xi_{\max}) \\ 0 & (\text{otherwise}) \end{cases} \tag{45.39}$$

where $g[\phi_2(\xi, \eta)] = \phi_2(\xi, \eta) \cos\beta_R \cos\psi_1 + \sin \sin\psi_1 \sin[\phi_2(\xi, \eta)]$; integrating the above equation yields the energy density at the receiving point in (J/m^2):

$$H_p = \int_{t_{\min}}^{t_{\max}} E(ct/r) dt. \tag{45.40}$$

The energy received by the receiver is

$$H = A_r . H_R, \tag{45.41}$$

where A_r is the receiving aperture area of the receiver; and then, the path loss is

$$L = 10\lg \frac{E_T}{H}, \tag{45.42}$$

where E_T is the pulse energy emitted by the transmitter.

Multiple-Scattering Model and Theory

The multiscattering propagation model of nonline-of-sight noncoplanar ultraviolet light is shown in Fig. 45.9. We set the transmitting end at the origin of the coordinate system (x, y, z) and the receiving end on the positive semiaxis of the x-axis, and the distance between Rx and Tx is d. Cones C_t and C_r represent the emitted beam and the receiving field of view, respectively. θ_t and ϕ_t represent the elevation and divergence angle half-angles of the transmitting end, respectively. θ_r and ϕ_r represent the receiving elevation and field-of-view angle half-angles, respectively.

α_t is the deflection angle of C_t and α_r is the deflection angle of C_r. S_n is the scattering point of the nth scatter. Here, r_0 is the distance from the transmitter to S_1, r_1 is the distance from the receiver to S_1, and r_n is the distance to S_n. ς_{sn} is the angle between the line of S_n and Rx on the central axis and is the angle between the

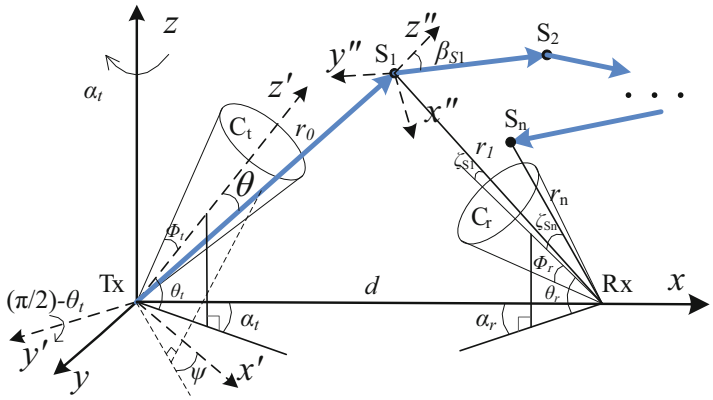


Fig. 45.9 Multiple-scattering propagation model of noncoplanar non-line-of-sight ultraviolet light

direction of photon incidence at the point and the direction of propagation after the photon scattering.

To conveniently describe the direction and distance of a photon transmission, coordinate system (x, y, z) needs to be transformed. The point coordinates based on coordinate system (x', y', z') and those based on coordinate system (x, y, z) are transformed according to Eq. (45.43):

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_{y'}\left(\frac{\pi}{2} - \theta_t\right) R_z(\alpha_t) \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{45.43}$$

where $R_{y'}\left(\frac{\pi}{2} - \theta_t\right) R_z(\alpha_t)$ is

$$R_{y'}\left(\frac{\pi}{2} - \theta_t\right) R_z(\alpha_t) = \begin{bmatrix} \cos\left(\frac{\pi}{2} - \theta_t\right)\cos\alpha_t & \cos\left(\frac{\pi}{2} - \theta_t\right)\sin\alpha_t & -\sin\left(\frac{\pi}{2} - \theta_t\right) \\ -\sin\alpha_t & \cos\alpha_t & 0 \\ \sin\left(\frac{\pi}{2} - \theta_t\right)\cos\alpha_t & \sin\left(\frac{\pi}{2} - \theta_t\right)\sin\alpha_t & \cos\left(\frac{\pi}{2} - \theta_t\right) \end{bmatrix}. \tag{45.44}$$

Suppose a photon is uniformly emitted in the emitting cone Tx, and any single photon is taken. In coordinate system (x', y', z'), the angle between the propagation direction of the photon and the forward direction of the coordinate axis z' is set to θ. The angle between the photon projection and the coordinate axis in the plane (x', y') is set to ϕ; (θ, ϕ) specifies the direction of the photon. θ is uniformly distributed among $(0, \phi_t)$ and can be found to be

$$\begin{aligned} \cos\theta &= 1 - \xi^{(\theta)}(1 - \cos\phi_t) \\ \phi &= 2\pi\xi^{(\phi)}, \end{aligned} \tag{45.45}$$

where $\xi^{(\theta)}$ and $\xi^{(\phi)}$ are random numbers that follow a uniform distribution in the [0, 1] interval. Thus, in coordinate system (x', y', z'), the directional cosine of the direction of the propagation of the emitted photons can be expressed as

$$\left(u_{x'},\ u_{y'},\ u_{z'}\right) = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta). \tag{45.46}$$

From Eq. (45.43), the directional cosine of the photon propagation direction in coordinate system (x, y, z) can be expressed as

$$\begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} = \left[R_y\left(\frac{\pi}{2} - \theta_t\right)R_z(\alpha_t)\right]^{-1} \begin{bmatrix} u_{x'} \\ u_{y'} \\ u_{z'} \end{bmatrix}. \tag{45.47}$$

In the photon transmission process, both Rayleigh and Mie scattering occur; therefore, the scattering phase function can be expressed as the weighted sum of Rayleigh and Mie scattering:

$$P(\cos\beta_s) = \frac{k_s^R}{k_s}P^R(\cos\beta_s) + \frac{k_s^M}{k_s}P^M(\cos\beta_s). \tag{45.48}$$

In the photon propagation process, the angle between the directions of propagation and incidence, that is, the angle of scattering, is determined by the scattering phase function:

$$\xi^{(S)} = 2\pi\int_{-1}^{\mu_{S1}} P(\mu)\mathrm{d}\mu, \tag{45.49}$$

where $\mu_{s1} = \cos\beta_{s1}$ and $\xi^{(S)}$ is a random variable that is uniformly distributed between 0 and 1. $P(\mu)$ can be derived from Eq. (45.48).

The probability that a photon will reach the receiving surface is

$$P_{1n} = \frac{A\cos\zeta_{sn}}{4\pi {r_n}^2}P(\cos\beta_{sn}), \tag{45.50}$$

where A is the area of the receiving aperture and $P(\cos\beta_{sn})$ is the phase function of the n^{th} scattering. The probability that a photon will travel a distance r_n after scattering by point S_n is

$$P_{2n} = e^{-k_e r_n}. \tag{45.51}$$

Therefore, the probability that a photon will reach the receiving surface after the n^{th} scattering is

$$P_n = W_n P_{1n} P_{2n}, \tag{45.52}$$

where W_n is the probability of survival before the photon arrives. W_n can be expressed as

$$W_n = (1 - P_{n-1})\mathrm{e}^{-k_a|\mathrm{S}_n - \mathrm{S}_{n-1}|}W_{n-1}, \tag{45.53}$$

where $|\mathrm{S}_n\text{- }\mathrm{S}_{n\text{-}1}|$ is the distance from scattering point $\mathrm{S}_{n\text{-}1}$ to scattering point S_n.

The probability $P_0 = 0$ of photon scattering directly reaching the receiving end is 0. As the amount of photon scattering increases, the probability of photon survival is smaller, so the maximum photon scattering number is set to $N \leq 5$. During the photon motion, each scattering that occurs is independent of the others, and the probability of a photon arriving at S_n in the receiving plane is

$$P_N = \sum_{n=1}^{N} P_n. \tag{45.54}$$

Suppose that the emitter emits M photons. In this study, $M = 10^6$. The total probability that photon m can travel to the receiving surface of the receiver by a maximum of N scattering times is $(P_N)_m$. The average probability that a photon from the emitter will reach the receiving surface is

$$P = \frac{\sum_{m=1}^{M} (P_N)_m}{M}, \tag{45.55}$$

where $(P_N)_m$ can be obtained using (45.54). P reflects the average probability that a photon from the emitting end reaches the receiving end. The emitting end emits M independent photons. Then, the number of photons that can reach the receiving end is $M \times P$. The path loss of a nonline-of-sight ultraviolet communication system can be calculated as

$$PL = 10\log_{10}\left(\frac{M}{M \times P}\right) = 10\log_{10}\left(\frac{1}{P}\right). \tag{45.56}$$

Path Loss Analysis of Ultraviolet Optical Communications

As shown in Fig. 45.10, there are three NLOS communication modes [67]:

Mode 1: Vertical transceiver mode ($\beta_T = \beta_R = 90°$)
Mode 2: Vertical receive, oblique transmit or vertical transmit, oblique receive mode ($\beta_R = 90°, \beta_T < 90°$ OR $\beta_T = 90°, \beta_R < 90°$)
Mode 3: Oblique transmit–receive mode ($\beta_R < 90°, \beta_T < 90°$)

Assuming that the emitting and receiving optical axes are in the same plane and UV light is scattered once to the receiving plane, the three modes are simulated. The parameters are set as follows: The half-angle of the emitting side is $\theta_T = 45°$, the

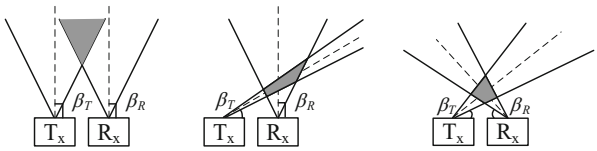


Fig. 45.10 Three modes of operation

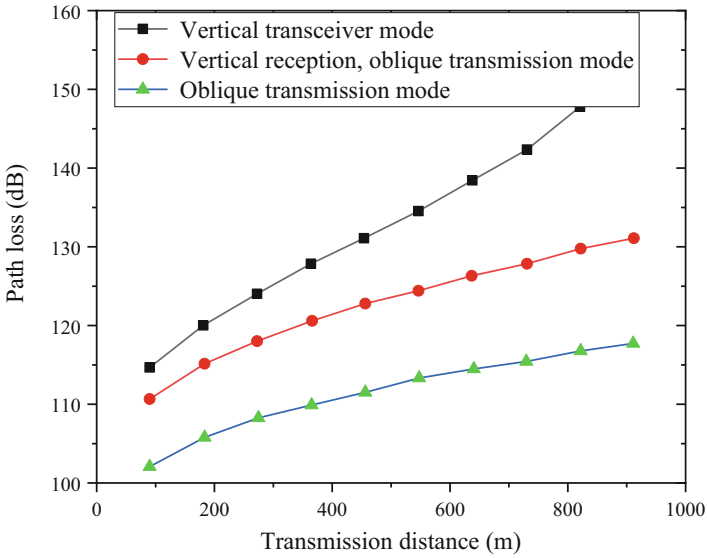


Fig. 45.11 Influence of the three working modes on the path loss

half-angle of the receiving side is $\theta_R = 30°$, wavelength $\lambda = 260$nm, Rayleigh scattering coefficient $\kappa_s^{\text{Ray}} = 0.24\text{km}^{-1}$, Mie scattering coefficient $\kappa_s^{\text{Mie}} = 0.25\text{km}^{-1}$, atmospheric absorption coefficient $\kappa_a = 0.9\text{km}^{-1}$, asymmetry factor $g = 0.72$, model parameter in the scattering phase function $r = 0.017$, and $f = 0.5$.

The simulation results are shown in Fig. 45.11. At the same transmission distance, the path loss of the vertical transceiver mode was the largest, followed by the vertical-receive and oblique-transmit mode, and the path loss of the oblique transmit and receive mode was the smallest. Therefore, when designing a transceiver system, the oblique transmit and receive mode should be chosen, which is more advantageous for ultraviolet communication under the same conditions.

In the third mode, the path loss differs for different receiving and transmitting elevation angles, and the parameter settings do not change. The divergence half-angle of the transmitter was set to $\theta_T = 10°$, and the field-of-view half-angle of the receiver was set to $\theta_R = 30°$. When the receiver angle was fixed at $\beta_R = 45°$, the elevation angle of the transmitter was changed to obtain the path loss. When the angle of the transmitter was fixed at $\beta_T = 45°$, the path loss was simulated by

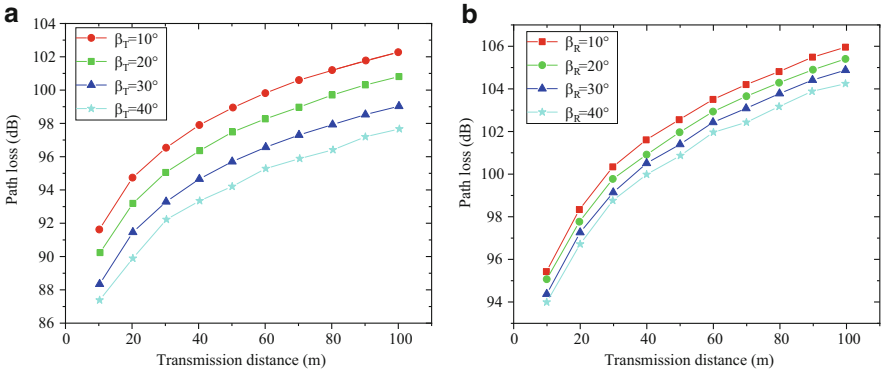


Fig. 45.12 Path loss for different transmit and receive elevation angles. (**a**) Different transmit elevations. (**b**) Different receive elevations

changing the elevation angle of the receiver. The simulation results are shown in Fig. 45.12.

The graph shows that the path loss increases with the increase of the transmit-receive angle, and the effect of changing the elevation angle of the transmitter is obviously greater than that of the receiver. The path loss increases with an increase in the elevation angle of the receiver, but the increase is small; therefore, the elevation angle of the receiver should be fixed and the elevation angle of the transmitter reduced to reduce the path loss.

In an ultraviolet communication system, the nonline-of-sight noncoplanar single-scattering model is more ideal than the NLOS noncoplanar multiple-scattering model. The path loss of the latter can be calculated using both the ergodic element method and the pointing probability method.

The receiving surface radius of the receiving end is 1.5×10^{-2} m, the distance between the receiving end and the sending end is $d = 100$ m, the angle of elevation of the sending end $\beta_T = 30°$, the angle of elevation of the receiving end $\beta_R = 60°$, and the angle of divergence of the sending end $\theta_T = 10°$. The angle of the field of view of the receiving end $\theta_R = 20°$, the angle of offset of the sending end $\alpha_t = 0°$, and the angles of offset of the receiving end are $\alpha_r = \{0°, 5°, 10°, 15°, 20°, 25°, 30°, 35°, 40°, 45°, 50°\}$.

In the pointing probability method, the number of photons sent is set to 10^6 and the parameters of the scattering phase function (γ, g, f) are set to (0.017, 0.72, 0.5). The atmospheric absorption coefficients are the Rayleigh scattering coefficient $k_S^{\text{Ray}} = 0.24 \times 10^{-3}\text{m}^{-1}$ and Mie scattering coefficient $k_S^{\text{Mie}} = 0.25 \times 10^{-3}\text{m}^{-1}$. The simulation results are shown in Fig. 45.13.

Figure 45.13 shows that the path loss increases and the volume of the common scatterer decreases with an increase in the offset angle at the end. The path loss calculated using the pointing probability method and the path loss energy calculated using the ergodic-microelement method both fit well.

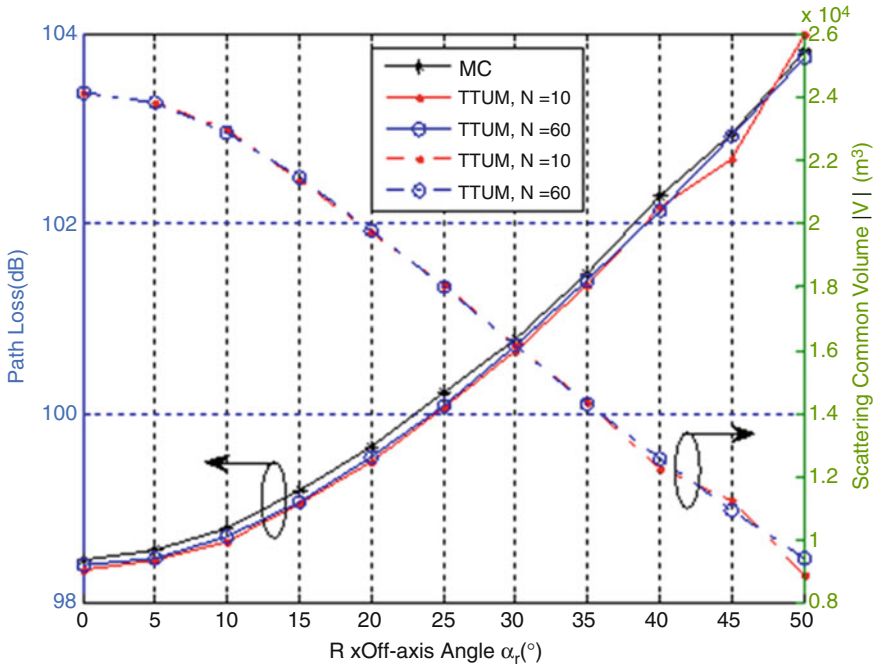


Fig. 45.13 Relationship between path loss and volume of the common scatterer and offset angle of the receiver [68]

UV Light Scattering Channel Characteristics Under Different Weather Conditions

UV Light Scattering Channel Characteristics on Rainy Days

When the ultraviolet light signal and raindrop particles are scattered in a rainy environment, the received ultraviolet light signal will be attenuated to varying degrees. In severe cases, it will interrupt the ultraviolet optical communication link. Therefore, to ensure the normal operation of the communication link, it is necessary to study the attenuation effect of raindrops of different sizes on ultraviolet light. At present, the main models used to describe the shape of raindrop particles are the approximate ellipsoid model, the Beard–Chuang raindrop-shape model, and the Pruppacher–Pitter shape model [69].

When the radii of the raindrop particles are greater than 0.5 mm, the Pruppacher–Pitter approximate raindrop-shape model should be used. The model uses the ratio of the major and minor axis radii a_r/b_r, which can approximate the degree of deformation of raindrop particles. A schematic diagram of the ellipsoid shape is shown in Fig. 45.14. The ellipsoid is formed by the rotation of its major axis a_r and minor axis b_r, and the corresponding ellipsoid shape is described as follows [70]:

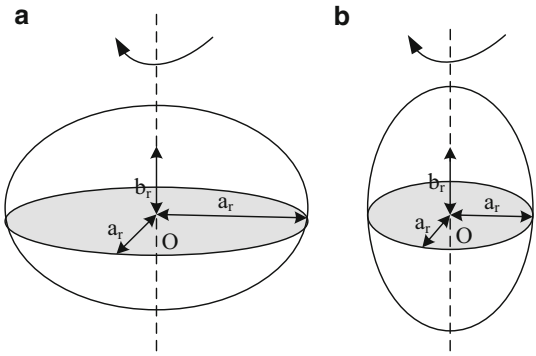


Fig. 45.14 Rotationally symmetric ellipsoids. (**a**) Oblate ellipsoid. (**b**) Long ellipsoid

$$r(\theta_p) = a_r\left(\sin^2\theta_p + \frac{a_r^2}{b_r^2}\cos^2\theta_p\right)^{-1/2}, \tag{45.57}$$

where θ_p represents the angle between the rotation angle and the longitudinal axis.

For deformed raindrop particles, the following mathematical expression [70] applies the Beard–Chuang (BC) model to the generic Chebyshev particle to describe the shape of the particle:

$$r(\theta_p) = r_0\left[1 + \sum_{n_p=0}^{10} c_n \cos(\theta_p n_p)\right], \tag{45.58}$$

where $r(\theta_p)$ is the particle radius corresponding to different orientation angles θ_p; r_0 is the radius of an equal area sphere. c_n is a constant, and $\cos(\theta_p n_p) = T_n(\cos\theta_p)$, which represent Chebyshev polynomials of the n_p order.

The BC model for calculating the shape of raindrops can be approximated as an ellipsoid, based on the deformation of raindrop particles while falling; this is called the approximate ellipsoid model. By considering the symmetry of the cross-section of the raindrop model, the area of the cross-section of the BC model was approximated by volume equality. Suppose the major axis of the approximate ellipsoid model section is a_{pr}; the minor axis is c_{pr}. Since the ellipsoid section is symmetrical along the abscissa axis and the ordinate axis, the short-axis c_{pr} of the approximate ellipsoid is solved by taking half the length of the corresponding longitudinal axis in the BC model, which can be expressed as [70]

$$c_{pr} = \frac{1}{2}\left[r(0^\circ) + r(180^\circ)\right] = r_v(1 + c_0 + c_2 + c_4 + c_6 + c_8 + c_{10}). \tag{45.59}$$

As the area of the section of the ellipsoid model is approximately equal to that of the BC model, the area of the vertical section of the BC model is expressed as

$$\begin{aligned} A &= r_v^2 \int_0^{\pi} \left[1 + \sum_{n_p}^{10} c_n \cos\left(n_p \theta_p\right) \right]^2 \\ &= \pi r_v^2 \left[1 + 2c_0 + c_0^2 + \sum_{n_p=1}^{10} \frac{c_n^2}{2} \right]. \end{aligned} \tag{45.60}$$

a_{pr} approximates the long axis of an ellipsoid model, which can be expressed as

$$a_{pr} = (1.05 - 0.131 r_v)^{-\frac{1}{3}} r_v. \tag{45.61}$$

Because raindrop particles in rainy environments have different scales, it is necessary to consider raindrop particles of different sizes in combination with the particle size distribution. A large amount of measured data shows that the size distribution of raindrop particles can be based on a negative exponential Marshall–Palmer (M-P) raindrop spectral distribution model. Its expression is

$$N(D) = 8 \times 10^3 \exp\left(-4.1 R^{-0.21} D\right), \tag{45.62}$$

where $N(D)$ represents the unit volume, that is, the number of raindrop particles in the unit radius interval, R represents the rainfall rate (mm/h), and D represents the radius of the raindrop particles.

The simulation results are shown in Fig. 45.15. It can be seen from the figure that for the same raindrop particle radius, the larger the rainfall rate, the more raindrop particles. The number of raindrops is distributed negatively exponentially with the increase of raindrop particle size.

In practical applications, raindrop particles of different sizes and number densities are randomly distributed in the atmosphere; therefore, the scattering effect of raindrop particles should be considered. The M–P spectral distribution can be used to calculate the attenuation caused by the raindrop particle swarm.

According to the relationship between the droplet size distribution and particle number density in Fig. 45.15, it can be concluded that the raindrop particle radius is mainly concentrated in the range of 0.1 mm–3 mm. The number of particles per unit volume can be determined based on the rainfall intensity. Using the above experimental conditions, a UV light path with a 260 nm wavelength under different rainfall conditions was simulated and analyzed.

The experimental wavelength $\lambda = 260$nm, elevation angles $\beta_R = \beta_T = 10°$, divergence angle $\theta_T = 30°$, receiving field angle $\theta_R = 60°$, receiving aperture $A = 1.66\ \text{cm}^3$, and a 1-mW emitting light power were assumed. The precipitation particle sizes, scattering coefficients, and absorption coefficients under different rainfall conditions are listed in Table 45.2.

Mie scattering is the main consideration in the scattering of UV light by raindrops. The results of the receiver power simulation are shown in Fig. 45.16.

Figure 45.16 shows that under different rainfall conditions, the power of the receiving end decreases with an increase in the communication distance, showing a

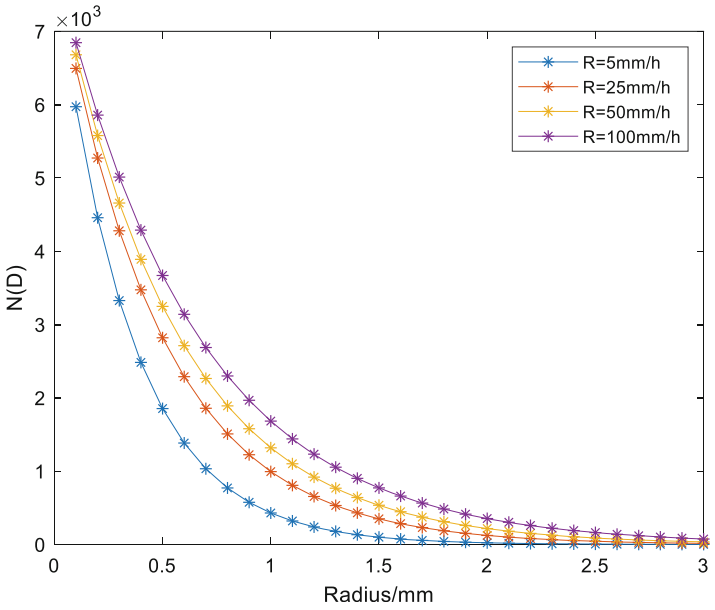


Fig. 45.15 Size distribution of raindrop particles

Table 45.2 Attenuation coefficients at different rainfall rates

Type of rainfall	Particle size (mm)	Scattering coefficient (m^{-1})	Absorption coefficient (m^{-1})
Drizzle	$0.1 < D < 1.5$	1.2×10^{-3}	0.906×10^{-3}
Widespread	$0.5 < D < 2$	3.4×10^{-3}	2.3×10^{-3}
Thunderstorm	$1 < D < 3$	9.3×10^{-3}	6.03×10^{-3}

negative index reduction trend. For the same communication distance, the transmission attenuation of ultraviolet light by thunderstorms is the largest, followed by widespread rain, and a drizzle has the least impact on it. This is mainly because the raindrop particles in a thunderstorm have a strong attenuation effect on ultraviolet light, so the received optical power is small.

UV Light Scattering Channel Characteristics on Foggy Days

According to the standards of the World Meteorological Organization, weather with a visibility of less than 1 km is attributed to foggy days; when the visibility is between 1 km and 10 km, it is considered light fog or haze. Weather with a horizontal visibility of less than 10 km is usually called haze. Haze is different from fog; its water vapor content is small, and the relative humidity is low, while the water vapor content in fog is basically saturated. The differences between foggy and hazy days are presented in Table 45.3.

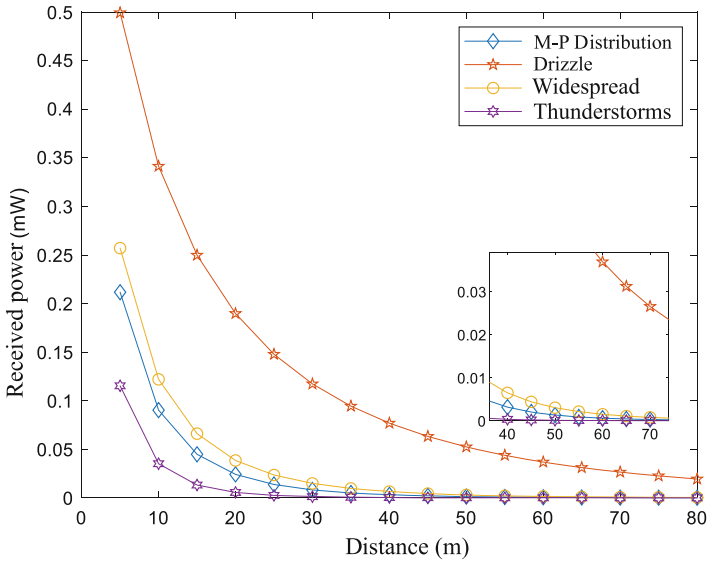


Fig. 45.16 Received power loss under different rainfall conditions

Table 45.3 Foggy and hazy weather characteristics [71]

Weather phenomenon	Composition	Relative humidity	Particle size	Visibility	Color	Concentration distribution
Foggy days	Tiny water droplets, ice crystals	$\geq 90\%$	1 μm–10 μm	< 1 km	Sky is milky white or cyan	Uneven
Hazy days	Fine, dry dust particles, carbon-containing particles, and soluble salts	$\leq 80\%$	0.001 μm–1 μm	< 10 km	Sky appears yellow or orange	Homogeneous and stable

The scale of rainfall and fog particles influences the scattering characteristics of ultraviolet light, and the scale parameters depend on the particle shape, particle size, and concentration distribution. Table 45.4 shows the radii and number densities of atmospheric particles.

Two models are commonly used for the distribution of fog particles: the modified gamma (Γ) function and the generalized gamma distribution. The parameters of the modified Γ function need to be determined by experimental data, so they have fitting errors. The generalized gamma distribution has wide applications, not only for fog particles over continents and oceans but also for aerosols in the stratosphere. The gamma-spectrum model is expressed as follows [72]:

Table 45.4 Radii and concentrations of atmospheric particles [71]

Type	Radius (μm)	Concentration (A/cm^3)
Air molecules	10^{-4}	10^{19}
Aitken	10^{-3}–10^{-2}	10^{-4}–10^{2}
Smog particles	10^{-2}–1	10–10^{3}
Fog particles	1–10	10–10^{2}
Rain particles	10^{2}–10^{4}	10^{-5}–10^{-2}

$$n(r_p) = a_d r_p^{\alpha_d} \exp\left(-b_d r_p^{\beta_d}\right), \tag{45.63}$$

where r_p is the radius of the fog particle, α_d and β_d are the gamma distribution index and are expressed as $a_d = (9.781/(V^6.\ W^5)) \times 10^{15}$, $b_d = (1.304/(V.\ W)) \times 10^4$, and V W are the visibility and liquid water content, respectively. Based on this spectral distribution model, the droplets are divided into advection droplets and radiation droplets, according to the empirical relationship between the water content and visibility [72]:

$$n(r) = 1.059 \times 10^7 V^{1.15} r^2.\exp\left(-0.8359 V^{0.43} r\right) \tag{45.64}$$

$$n(r) = 3.104 \times 10^{10} V^{1.7} r^2.\exp\left(-4.122 V^{0.54} r\right), \tag{45.65}$$

where V is the visibility in km, r is the radius of fog particles in μm, and $n(r)$ is in $\mathrm{m}^{-1}\mu\mathrm{m}^{-1}$.

Thick fog ($V = 0.05$ km), medium fog ($V = 0.2$ km), and mist ($V = 2$ km) were analyzed. From Eqs. (45.64) and (45.65), the size distributions of the advection and radiation droplets were obtained, as shown in Fig. 45.17.

As shown in Fig. 45.17, at the same visibility, the diameter of advection droplets is larger than that of radiation droplets at the highest droplet concentration: the concentration of advection droplets is two orders of magnitude lower than that of radiation droplets. The lower the visibility, the larger the radius of the droplet under the highest droplet concentration. With a decrease in visibility, the maximum droplet concentration gradually decreases.

When analyzing the performance of direct ultraviolet optical communication systems, the attenuation of optical signals in the atmospheric channel has an important impact on communication performance, and accurately and easily analyzing ultraviolet light attenuation is key. In foggy weather, light attenuation caused by fog is often analyzed using empirical attenuation models, such as the Kim, Kruse, and Nabousi models [72].

The empirical formula used by the Kim and Kruse models is [73]

$$\alpha = \frac{3.912}{V}\left(\frac{0.55}{\lambda}\right)^q, \tag{45.66}$$

where V is the visibility in km, λ is the wavelength of the incident light in μm, q is a constant related to V, and the q values of the Kim and Kruse models are different, as follows:

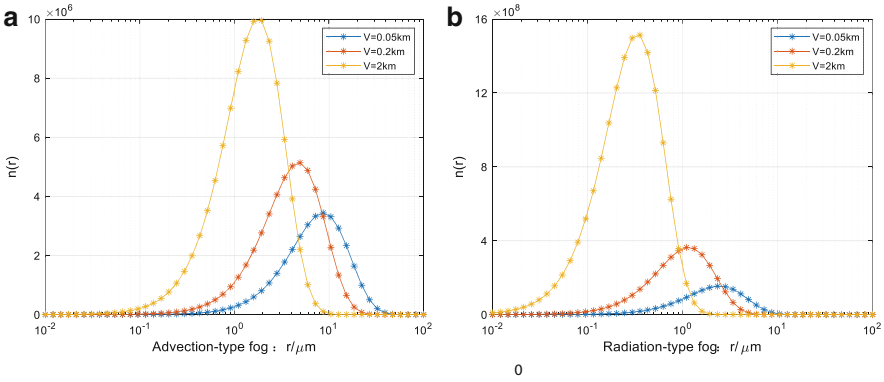


Fig. 45.17 Size distribution of different types of fog particles. (**a**) Advection-type fog particles. (**b**) Radiation-type fog particles

$$q_{\text{Kim}} = \begin{cases} 1.6 & \text{V} > 50\text{km} \\ 1.3 & 6\text{km} < \text{V} < 50\text{km} \\ 1.6V + 0.34 & \text{V} < 6\text{km} \\ V - 0.5 & 0.5\text{km} < \text{V} < 1\text{km} \\ 0 & \text{V} < 0.5\text{km} \end{cases} \tag{45.67}$$

$$q_{\text{Kruse}} = \begin{cases} 1.6 & V > 50\text{km} \\ 1.3 & 6\text{km} < V 50\text{km} \\ 0.585V^{\frac{1}{3}} & V < 6\text{km} \end{cases} \tag{45.68}$$

The Nabousi model divides fog into advection and radiation fogs, and its attenuation models are [74] the following:

$$\alpha_{\text{adv}} = \frac{0.11478\lambda + 3.8367}{V} \tag{45.69}$$

$$\alpha_{\text{rad}} = \frac{0.18126\lambda^2 + 0.13709\lambda + 3.7502}{V}, \tag{45.70}$$

where V is the visibility in km and λ is the wavelength of the incident light in μm. A simulation of the decay coefficients of the three empirical equations at an incident wavelength of 260 nm is shown in Fig. 45.18.

Overall, all three models exhibited a gradual decrease in attenuation with increasing visibility. As shown in Fig. 45.18, the attenuation curves of the Kim and Kruse models are almost identical at a 2-km visibility, while the magnitudes of the attenuation coefficients of the Nabousi model and the first two are quite different. However, as the visibility increased, the difference between the two models gradually decreased.

UV Light Scattering Channel Characteristics on Dusty Days

Sandy and dusty weather is caused by strong winds that carry sand, soot particles, etc., into the atmosphere. They are suspended in the atmosphere to form stable forms

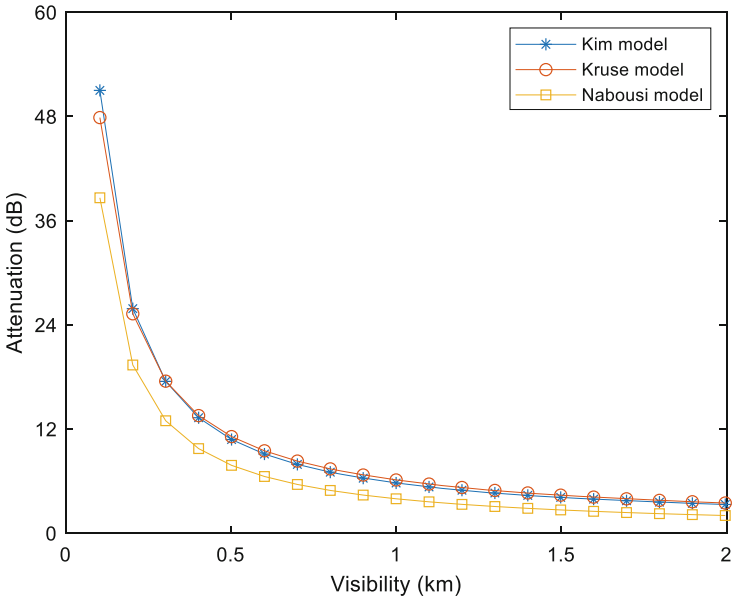


Fig. 45.18 Comparison of three empirical decay models

of dust particles, soot particles, aerosols, etc., resulting in low atmospheric visibility and decreased air quality. The scale spectrum distribution of dusty aerosol particles in the air differs in different grades of dusty weather. Aerosol particles in dusty weather include particles that are difficult to dissolve in water, such as sand, dust, and soot. The size distribution of aerosol particles in dusty weather follows a lognormal distribution, and the scale spectral distribution function is [75]

$$n(r) = \frac{N_0}{r\sqrt{2\pi}\lg\sigma} \exp\left\{-\frac{1}{2}\left[\frac{\lg r - \lg r_m}{\lg\sigma}\right]^2\right\}, \tag{45.71}$$

where r is the radius of the aerosol particle, r_m is the average radius of the particle, σ is the standard deviation, and N_0 is the particle number density. The scale spectral distribution probability density function of aerosols in this weather is expressed as

$$p(r) = \frac{1}{r\sqrt{2\pi}\lg\sigma} \exp\left\{-\frac{1}{2}\left[\frac{\lg r - \lg r_m}{\lg\sigma}\right]^2\right\}. \tag{45.72}$$

Table 45.5 shows the parameters of different types of aerosol commonly found in dusty weather [76].

Because N_0 is a very difficult physical quantity to measure, the concentration of dust aerosols is often described using the visibility, V, when studying the attenuation of UV light transmission during dusty weather:

Table 45.5 Scale distribution model parameters of different types of aerosol

	Dust type	Soot type
r_m	0.5	0.0118
σ	2.99	2.00

$$N_0 = \frac{15}{8.686 \times 10^3 \pi V \int_{r_1}^{r_2} r^2 p(r) dr}. \tag{45.73}$$

According to the parameters of different types of aerosol, combined with the lognormal distribution of the aerosol scale spectrum distribution function, the scale spectrum distribution of aerosol particles in different visibility and dusty weather conditions was obtained.

As shown in Fig. 45.19, the particle radii of different aerosol types differ in dusty weather. The particle radius of dust-type aerosols is mostly around 10^{-2} μm–1 μm while that of soot-type aerosols is mostly around 10^{-3} μm–10^{-1} μm.

Table 45.6 shows the particle number densities of dust-type and soot-type particles at different visibilities. As shown in the table, the higher the visibility, the lower the number density of the aerosol particles. Equation (45.72) shows that the probability density function of the same aerosol particle's scale spectrum distribution is the same. Under the same distribution probability density function, visibility is the only variable that affects the number and density of aerosol particles; therefore, visibility is inversely proportional to the number and density of aerosol particles.

In dusty weather, visibility affects the particle number concentration of aerosol particles and thus the light attenuation coefficient; therefore, this study considers visibility as the independent variable of the empirical attenuation model in dusty weather. From the Mie scattering theory, it can be seen that the attenuation ability of ultraviolet light through a unit distance is

$$A = 4.343 \times 10^3 \int_{r_1}^{r_2} N_0 C_{\text{ext}}(m, x) P(r) dr, \tag{45.74}$$

where $x = 2\pi r/\lambda$ is the aerosol particle scale parameter and the unit of A is dB/km. The above equation can be simplified to

$$\begin{aligned} A &= 4.343 \times 10^3 \int_{r_1}^{r_2} N_0 C_{\text{ext}}(m, x) P(r) dr \\ &= \frac{15 \int_{r_1}^{r_2} C_{\text{ext}}(r) p(r) dr}{2\pi V \int_{r_1}^{r_2} r^2 p(r) dr}. \end{aligned} \tag{45.75}$$

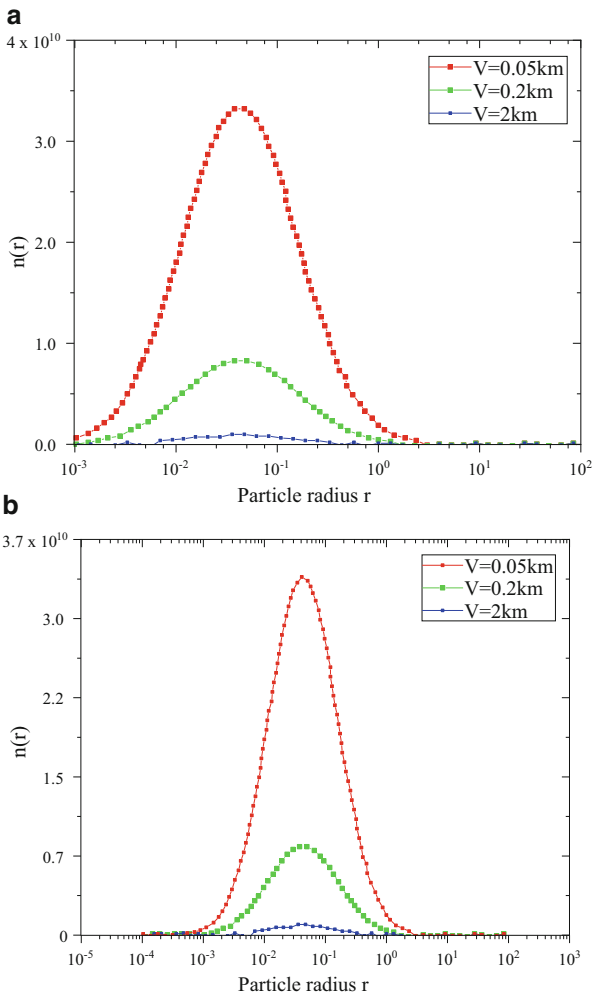


Fig. 45.19 Distribution of different particle scale spectra in dusty weather. (**a**) Dust-type particles. (**b**) Soot-type particles

Table 45.6 Number density of aerosol particles in different visibilities

		Particle number density	
Visibility/km	0.05	0.2	2
Dust-type particles	1.1×10^7	2.75×10^6	2.64×10^5
Soot-type particles	1.1×10^{10}	2.75×10^9	2.67×10^8

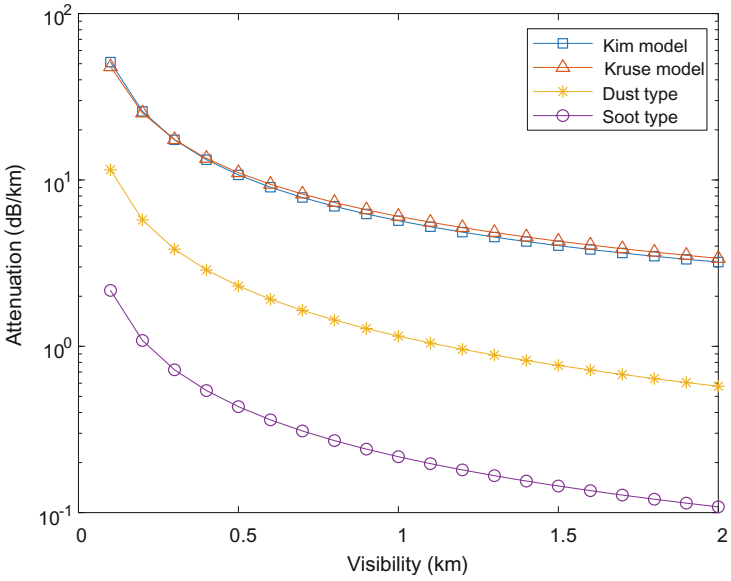


Fig. 45.20 Attenuation coefficients for four attenuation models

For a 260-nm incident wavelength, the results of simulating the attenuation coefficient using the above formula in dusty weather and the attenuation coefficient in foggy weather are shown in Fig. 45.20.

As shown in Fig. 45.20, in the ultraviolet light band, light attenuation in the two types of aerosol environment on foggy days is much greater than that in the two types of aerosol environments in dusty weather; both show exponential attenuation. Among the three aerosol types, the attenuation capacity of ultraviolet light from highest to lowest is fog, sand and dust, and soot.

Experimental Study of the Noise Model Under Different Weather Conditions

Experimental Setup

The experimental device shown in Fig. 45.21 was built to measure the ultraviolet background noise from sunrise to sunset on sunny days during the spring equinox and winter solstice seasons in Xi'an. In the figure, sunlight is transmitted through the atmospheric channel, reaching the surface of detector B (model OPHIR-PD 300-UV) through filter A (220 nm and 300 nm) on the detector surface. The detector then converts the optical signal into an electrical signal and transmits it to the computer. The StarLab 3.62 software was used to store and analyze the data. The

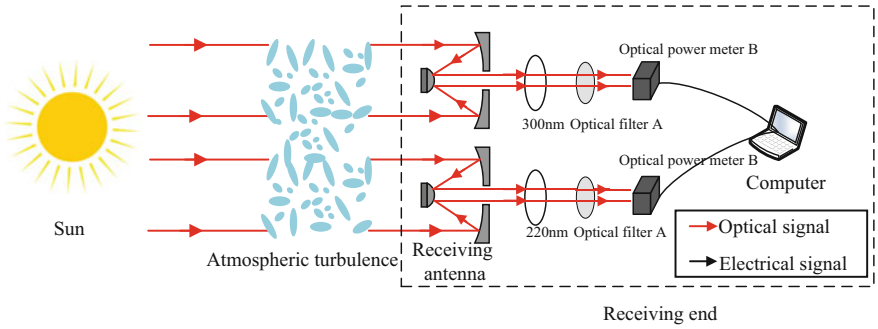


Fig. 45.21 Schematic diagram of experimental measurements

detector sampling frequency was 15 Hz, and the response wavelength was set to the 220-nm and 300-nm bands. The accuracy of the theoretical model was verified by comparison with measured data.

Near Earth Background Light Noise Model

When sunlight passes through the atmosphere, it is scattered by scattering elements, such as atmospheric molecules and aerosol particles, forming a background light in the sky. The complexity of the atmospheric channel constantly changes the radiant intensity of the sky background light, which increases the difficulty of theoretical calculations. The measured and theoretical values vary depending on the weather conditions.

Measured Data Analysis on Sunny Days

To intuitively observe the peak changes on different dates, the accuracy of the verification theoretical model (21) was compared, and the measured and theoretical values of the background noise of ultraviolet light during different periods of sunny days in January and February 2023 in Xi'an were compared. The measurement conditions matched the above simulation conditions, and the ultraviolet background noise in the 220-nm ±10-nm and 300-nm ±10-nm bands were measured simultaneously.

The results are shown in Fig. 45.22. The peak ultraviolet radiation on January 7 is small and the peak is about 12:41; with the passage of time, the daylight time becomes longer, the ultraviolet background radiation increases, the peak gradually increases, and the time of appearance is closer to 13:00. The peak offset reaches the maximum on February 27, and the peak size continues to increase. In sunny weather, the overall ultraviolet radiation trend was consistent with that of the theoretical model.

As shown in Fig. 45.22, the intensity of the UV radiation varies with different dates and bands. On January 7, the daily maximum radiation value at 300 nm was

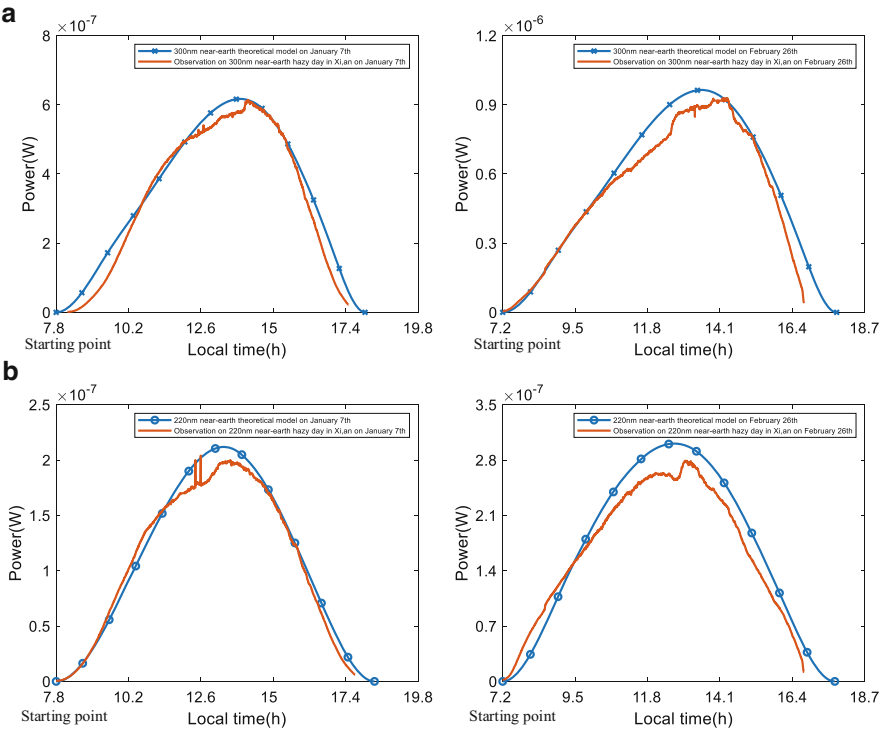


Fig. 45.22 Comparison between theoretical and measured results on a sunny day. (**a**) Results measured at different times on sunny days at a 300-nm wavelength. (**b**) Results measured at different times on sunny days at a 220-nm wavelength

6.14×10^{-7} W and the daily average UV radiation value was 3.46×10^{-7} W. On February 26, the daily maximum radiation at 300 nm was 1.03×10^{-6} W and the daily average ultraviolet radiation was 5.95×10^{-7} W. The maximum measured on February 26 in the 300-nm band was 1.68 times higher than the maximum on January 7.

On January 7, the daily maximum radiation at 220 nm was 2.04×10^{-7} W and the daily average ultraviolet radiation was 1.13×10^{-7} W. On February 26, the daily maximum radiation at 220 nm was 33.10×10^{-7} W and the daily mean UV radiation was 1.87×10^{-7} W. The maximum measured at 220 nm on February 26 was 1.52 times greater than the maximum on January 7.

Measured Data Analysis on Cloudy Days

The theoretical and measured values in sunny weather are basically the same; however, in cloudy weather, owing to the increase in cloud cover, the intensity of ultraviolet radiation is seriously attenuated. Moreover, owing to the influence of atmospheric turbulence and the movement of clouds, the fluctuation of ultraviolet radiation is large and the change trend is irregular.

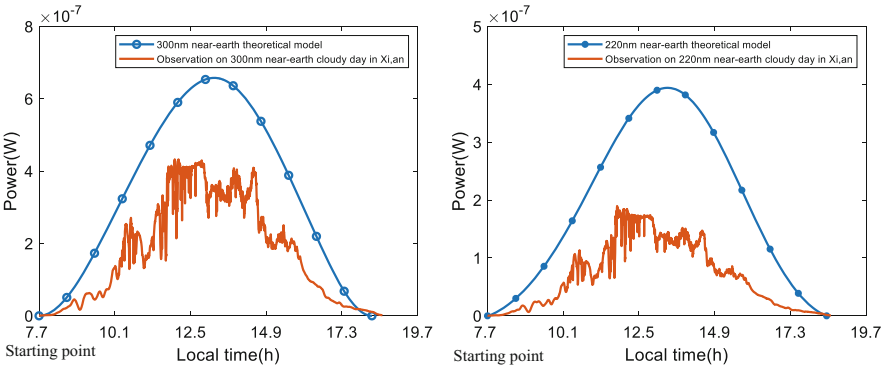


Fig. 45.23 Comparison between theoretical and measured data on cloudy days

As shown in Fig. 45.23, the actual measured values are generally in line with the theoretical trend; however, their values fluctuate greatly and exhibit obvious attenuation. Under cloudy weather, the daily maximum radiation value at the 300-nm band is 4.18×10^{-7} W and the daily average ultraviolet radiation value is 2.48×10^{-7} W. The daily maximum radiation value at the 220-nm band is 1.90×10^{-7} W and the average daily UV radiation value is 6.84×10^{-8} W. This differs from the theoretical value and cannot be used to describe the changing trend of UV radiation in cloudy weather.

Measured Data Analysis on Rainy Days

On rainy days, clouds become denser, the scattering and absorption of water vapor make the attenuation more severe, and the UV radiation amplitude is smaller. Compared with the severe waveband on overcast days, the fluctuation range is gentler and concentrated on rainy days, and the influence of atmospheric turbulence on the cloud layer is not significant.

As shown in Fig. 45.24, the actual measurement is similar to the theoretical trend, reaching a maximum at noon; however, its magnitude varies greatly. In cloudy weather, the maximum daily radiation value of the measured 300-nm band is 1.55×10^{-7} W, and the daily average ultraviolet radiation value is 5.44×10^{-8} W. The maximum daily radiation value of the 220-nm band is 7.00×10^{-8} W, and the daily average ultraviolet radiation value is 2.56×10^{-8} W. The difference between the two is large, compared with the theoretical value, and cannot well characterize the ultraviolet radiation trend in cloudy weather.

Analysis of Measured Data on Hazy Days

The influence of the cloud layer and surface-suspended particles in hazy weather makes the change in the UV background radiation more drastic. The change in the UV background radiation is not correlated with the change in the solar altitude angle but is related to the degree of air turbidity. In the morning, the air is clearer and the UV radiation is higher. As the turbidity of the air increases, the UV radiation

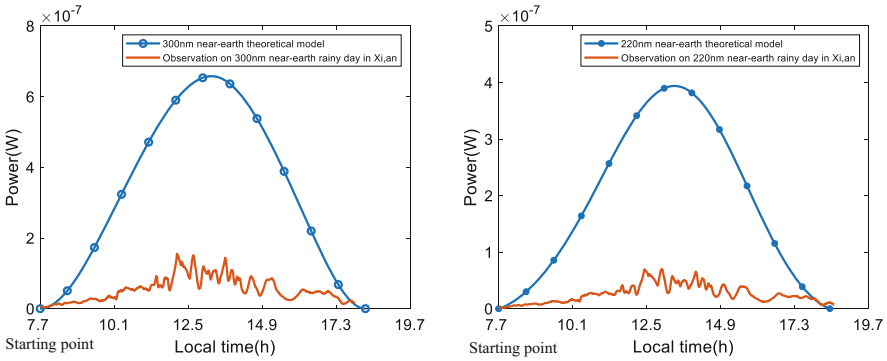


Fig. 45.24 Comparison between theoretical and measured data on rainy days

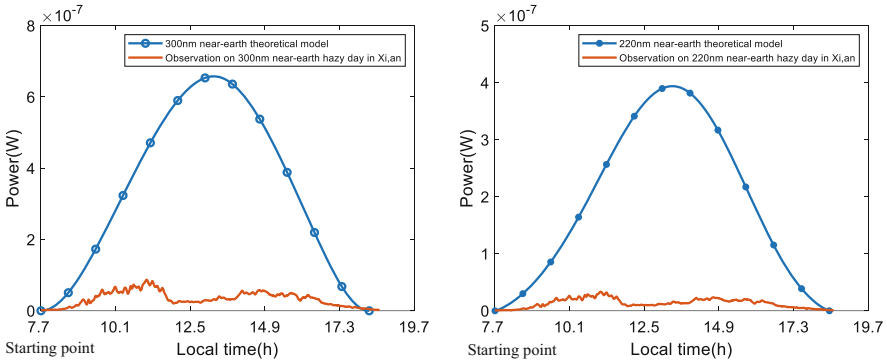


Fig. 45.25 Comparison between theoretical and measured data on hazy days

decreases continuously until the afternoon, when the UV radiation begins to rise; however, by this time, the Sun's altitude angle has begun to decrease. When the air turbidity recovers, there is a brief rise, and then the ultraviolet radiation begins to slowly decrease.

As shown in Fig. 45.25, the actual measured values do not conform to the theoretical trend, and the amplitude difference is large. Under hazy weather, the daily maximum radiation at 300 nm is 1.35×10^{-7} W and the daily average ultraviolet radiation is 4.44×10^{-8} W. The daily maximum radiation at 220 nm is 4.00×10^{-8} W and the daily mean UV radiation is 2.56×10^{-8} W. This is different from the theoretical value and cannot be used to describe the changing trend of UV radiation in hazy weather.

High-Altitude Background Light Noise Model

Based on an existing near-ground noise model, this study measured the ultraviolet background noise during part of a high-altitude period to restore the high-altitude

noise model. To ensure the accuracy of the high-altitude model, high-altitude ultraviolet background noise at 10 km in Urumqi, 8 km in Beijing, 8 km in Nanjing, and 7 km in Yantai were measured from Xi'an.

By comparing the measured data in the high-altitude time period with the near-ground noise model in this period, the expectation of an increased multiple of the high-altitude background noise was obtained, and the wave peak of the high-altitude noise model was determined. Cubic spline interpolation was used to select the correct data on both sides of the wave peak, according to the measured data and other time intervals, and the model was restored.

The principle is to divide the interval $[a, b]$ into n intervals $[((x_o, x_1), \ldots., (x_1, x_2))]$ and (x_{n-1}, x_n) with $n + 1$ points, two of which have two $x_o = a$, $x_n = b$ endpoints. Cubic spline interpolation uses the curve of each interval as a cubic equation. The cubic equation satisfies the condition that for each interval, $[x_i, x_{i+1}]$, $S(x) = S_i(x)$ is a cubic equation satisfying $S_i(x) = y_i$ $(i = 0, 1, \ldots, n)$ and $S(x)$, $S'(x)$ where $S''(x)$ is continuous. Then, these three conditions can be used to construct an equation $S_i(x) = a_i + b_i x + c_i x^2 + d_i x^3$ to obtain a_i, b_i, c_i, and d_i in a complete interpolation function by solving.

Xi'an–Urumqi 10-km Altitude

Figure 45.26 shows the noise model of the restoration of the 10-km high-altitude route between Xi'an and Urumqi. The roundtrip flights to Urumqi were HU7898 and GS7525 at 12:25–16:10 on March 11 and 7:45–11:15 on March 12, respectively.

As shown in the figure, the high-altitude noise model also conforms to the diurnal variation law of the solar altitude angle, in which the maximum value of the 300-nm high-altitude ultraviolet background noise model in Xi'an–Urumqi is 3.416 μW, and the peak appears at about 13:48, which is 3.322 times that of the same time near Earth. The 220-nm band is more seriously absorbed by ozone and the attenuation is more obvious. Its background noise maximum is 1.152 μW and the peak moment occurs at 12:48, which is 3.152 times that of the same time near Earth.

Xi'an–Beijing 8-km Altitude

Figure 45.27 shows the noise model of the 8-km high-altitude route between Xi'an and Beijing. The roundtrip flights to Beijing were HU7238 and HU7237 at 15:30–18:00 on March 6 and 7:50–10:15 on March 7, respectively.

As the figure shows, with a decrease in altitude, the propagation path of sunlight increases and the UV radiation decreases. The maximum value is 2.631 μW and the peak appears at about 13:36, which is 2.623 times of that at the same time on the surface. At 220 nm, the absorption of ozone is more serious and the attenuation is more obvious. The maximum value of background noise is 0.967 μW and the peak time is 12:36, which is 2.696 times that at the same time on the surface.

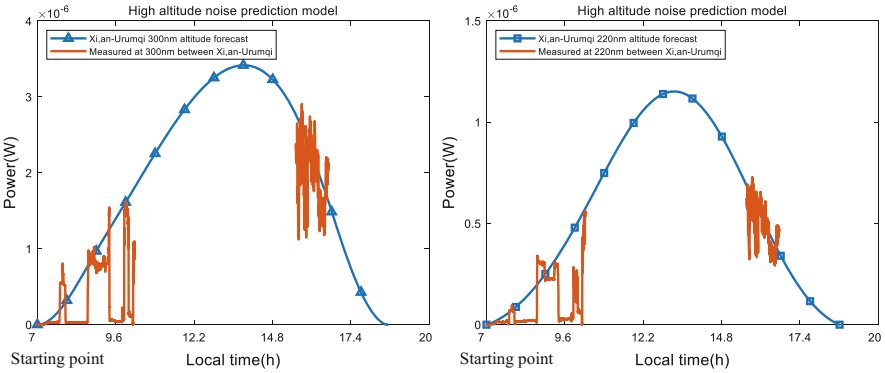


Fig. 45.26 10-km altitude between Xi'an and Urumqi on March 11: (left) 300 nm and (right) 220 nm

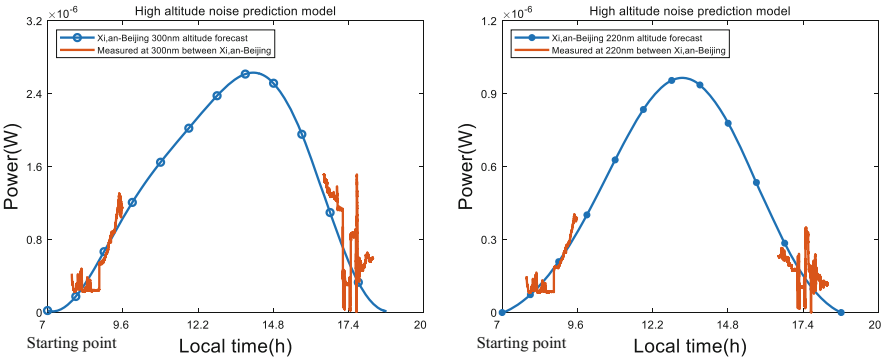


Fig. 45.27 Xi'an–Beijing 8-km altitude on March 7: (left) 300 nm and (right) 220 nm

Xi'an–Nanjing 8-Km Altitude

Figure 45.28 shows the noise model of the restoration of the 8-km high-altitude route between Xi'an and Nanjing. The roundtrip flights to Nanjing were 3 U3259 and MU2200 at 7:00–9:05 on April 15 and 15:55–17:50 on April 16, respectively.

As shown in the figure, the maximum of 300-nm UV background noise in Xi'an–Nanjing is 2.551 μW, which is close to the UV radiation at an 8-km altitude in Xi'an–Beijing, and the peak appears at about 13:42, 2.552 times as large as that of the surface at the same time. The 220-nm band is close to the Beijing route and the maximum background noise is 0.951 μW. The peak time appears at approximately 13:00, which is 2.58 times that at the same time on the surface.

Xi'an–Yantai 7-km Altitude

As shown in Fig. 45.29, the noise model of the Xi'an–Yantai 7-km high-altitude flight route was restored. Flights 9H8397 and HU7506 to Yantai ran from 15:50 to 17:50 on March 31 and from 9:40 to 12:00 on April 3, respectively.

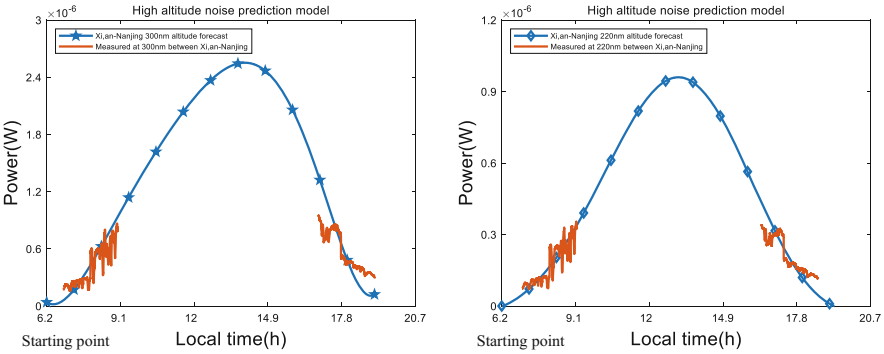


Fig. 45.28 Xi'an–Nanjing at an altitude of 8 km on April 15: (left) 300 nm and (right) 220 nm

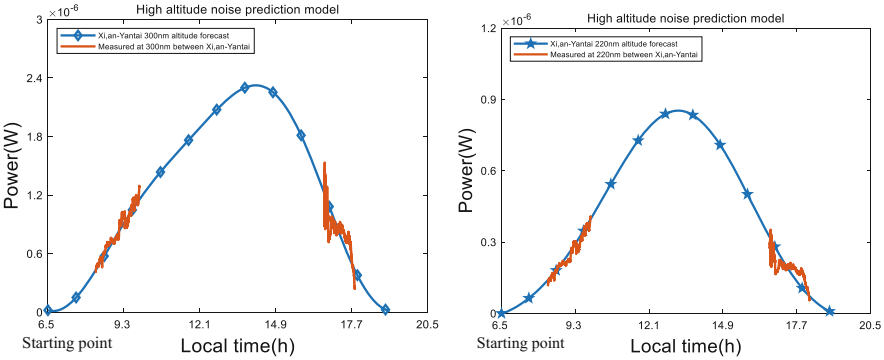


Fig. 45.29 Xi'an–Yantai at an altitude of 7 km on March 31. (left) 300 nm and (right) 220 nm

As shown in the figure, the background noise of 300-nm ultraviolet light at a high altitude from Xi'an to Yantai increases with a decrease in altitude; the propagation path of sunlight increases and the ultraviolet radiation decreases. The maximum value is 2.324 μW and the wave peak appears at about 14:00, 2.260 times that of the surface at the same time. The maximum value of UV background noise at 220 nm is 0.853 μW, and the peak time appears at about 13:00, which is 2.33 times of that at the same time on the surface.

Table 45.7 lists the peak moment, maximum peak, and multiple of the intensity of the UV background radiation compared to the near Earth UV background radiation under different high-altitude conditions.

The experimental results show that the UV background noise at 220 nm is consistent with the variation in the solar altitude angle and reaches its maximum near 13:00. However, the peak of the UV background noise at 300 nm is slightly shifted, and the peak appears at approximately 14:00. The maximum ultraviolet background noise values at 300 nm at 7-km, 8-km, and 10-km altitudes were 2.324 μW, 2.631 μW, and 3.416 μW, respectively, which were 2.26 times, 2.623

Table 45.7 Model parameters of high-altitude background noise in different cities

Region/wavelength	Altitude (km)	Peak time (h)	Peak amplitude (μW)	Multiple
Urumqi/300 nm	10	13.8	3.416	3.322
Urumqi/220 nm	10	12.8	1.152	3.152
Beijing/300 nm	8	13.6	2.631	2.623
Beijing/220 nm	8	12.6	0.967	2.696
Nanjing/300 nm	8	13.7	2.551	2.552
Nanjing/220 nm	8	13	0.951	2.58
Yantai/300 nm	7	14	2.324	2.26
Yantai/220 nm	7	13	0.853	2.33

times, and 3.322 times the surface noise, respectively. UV background noise values at 220 nm were 0.853 μW, 0.967 μW, and 1.152 μW, which were 2.33 times, 2.696 times, and 3.15 times that of the surface noise, respectively.

Above the clouds, UV radiation increases by 20–23% for every 1000-m increase in altitude. For an all-weather ultraviolet communication system, the threshold of the filter can be flexibly adjusted to filter the background noise according to the time change in clear weather to improve the stability of the UV communication system.

Data Correlation Function Analysis

As shown in Fig. 45.30, the correlation function of the 300-nm and 220-nm bands at high altitudes in each city shows that the similarity between the two bands is relatively high. This indicates that the measurement environment of the two bands is the same, the change trend is similar, and the measurement results are not affected by other factors. The similarity between the two bands is the strongest at the same time, and as the delay increases, the correlation coefficient decreases exponentially, indicating an uncorrelated nature.

Analysis of Light Intensity Distribution Characteristics of Ultraviolet Optical Communication

Ultraviolet Indirect Turbulence Model

Figure 45.31 shows a geometric sketch of an NLOS UV single-scattering transmission when there is atmospheric turbulence in the UV communication link. The NLOS transmission path is divided into two LOS links: the first from the Tx to the common volume V (link length l) and the other from the scattering volume V to the Rx (link length L). It is assumed that the Tx emission plane is C_t and the Tx beam angle is small, so that the flicker caused by turbulence in the scattering volume V is constant. To facilitate processing, the received optical power is calculated using the single-scattering principle.

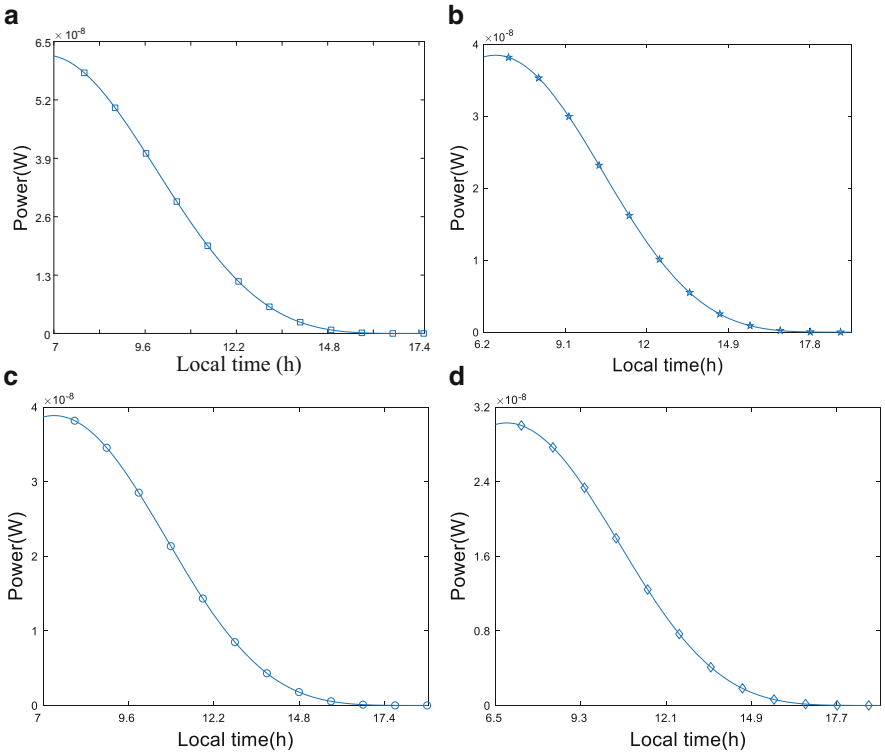


Fig. 45.30 High-altitude 300-nm and 220-nm-band correlation functions. (**a**) Correlation function for Xi'an–Urumqi. (**b**) Correlation function for Xi'an–Beijing. (**c**) Correlation function in Xi'an–Nanjing. (**d**) Correlation function in Xi'an–Yantai

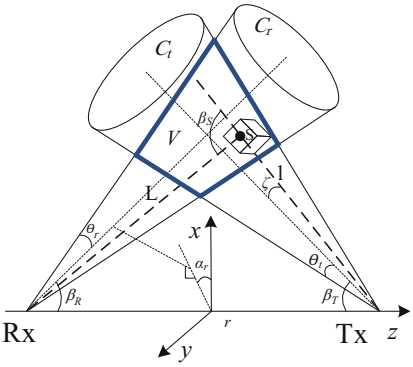


Fig. 45.31 Ultraviolet optical communication link in a turbulent atmosphere

When considering UV attenuation, both the optical power attenuation caused by scattering and absorption and the flicker attenuation caused by turbulence are considered to obtain a more accurate optical power PDF. The flickering attenuations of the two LOS links are

$$\alpha_{\text{turb_}l} = 2.\sqrt{23.17.C_n^2.(2\pi)^{7/6}.l^{11/6}} \tag{45.76}$$

$$\alpha_{\text{turb_}L} = 2.\sqrt{23.17.C_n^2.(2\pi)^{7/6}.L^{11/6}}. \tag{45.77}$$

The flicker attenuation caused by turbulence was simulated using Eq. (45.76), as shown in Fig. 45.32. It can be seen that as the communication distance or turbulence intensity increases, the flicker attenuation increases. In addition, it can be seen from the above equation that as the wavelength of the light wave decreases, the flicker attenuation increases and the flicker becomes more intense. This implies that UV light is more affected by turbulence than traditional free space optical communication, owing to its smaller wavelength.

The ultraviolet optical communication link is composed of links from the transmitting end to the scattering point and the scattering point to the receiving end. Then, the flicker attenuation of the entire indirect UV optical communication link is

$$\alpha_{\text{turb}} = \alpha_{\text{turb_}l} + \alpha_{\text{turb_}L}. \tag{45.78}$$

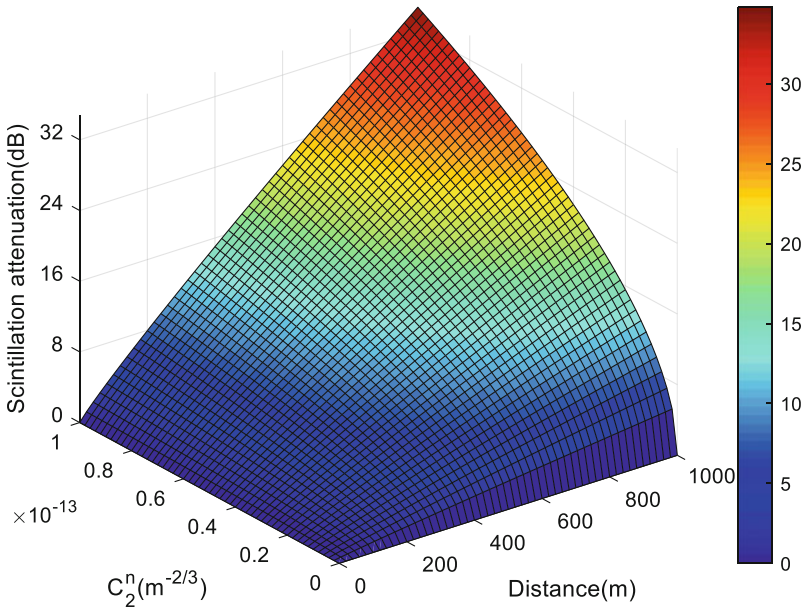


Fig. 45.32 Scintillation attenuation as a function of turbulence intensity and distance

For the link before scattering, the log normal distribution of the optical power reaching the scattering volume V PDF is

$$p(P_i) = \frac{1}{\sqrt{2\pi}\sigma_l P_l} \exp\left\{ -\frac{\left[\ln(P_l/P_{l0}) - \langle \ln(P_l/P_{l0})\rangle\right]^2}{2\sigma_l^2} \right\}, \tag{45.79}$$

where P_i is the power at scattering volume $V$and σ_l^2 is determined by a plane wave or spherical wave. The average optical power when the turbulence is not considered is $P_{l0} = P_T k_s V \exp(-k_s l)/(\Omega_T l^2)$, where P_T is the emitted optical power. $\langle P_l/P_{l0}\rangle$ can be expressed as

$$\langle P_l/P_{l0}\rangle = \exp\left[\left\langle \ln(P_l/P_{l0}) + \sigma_l^2/2\right\rangle\right]. \tag{45.80}$$

Defining α_{PL_l} as the path loss due to absorption and scattering,

$$\alpha_{PL_l} = 10\lg(P_T/P_{l0}). \tag{45.81}$$

When accounting for attenuation due to turbulent flickering, the total path loss is

$$\alpha_{\text{turb}_l} + \alpha_{\text{turb}_L} = 10\lg(P_T/\langle P_l\rangle). \tag{45.82}$$

A geometric transformation of the above equation can be obtained:

$$\langle \ln(P_l/P_{l0})\rangle = -\sigma_l^2/2 - \left(\alpha_{\text{turb}_l}.\ln 10\right)/10. \tag{45.83}$$

Substituting Eq. (45.82) into Eq. (45.79) yields

$$p(P_l) = \frac{\exp\left\{ -\frac{\left[\ln(P_l/P_{l0}) + \sigma_L^2/2 + \left(\alpha_{\text{turb}_l}\cdot \ln 10\right)/10\right]^2}{2\sigma_L^2} \right\}}{\sqrt{2\pi}\sigma_l P_l}. \tag{45.84}$$

Similarly, for the scattered link, the conditional PDF for the optical power received by Rx is

$$p(P_R, P_l) = \frac{\exp\left\{ -\frac{\left[\ln(P_R/P_{L0}) + \sigma_L^2/2 + \left(\alpha_{\text{turb}_L}\cdot \ln 10\right)/10\right]^2}{2\sigma_L^2} \right\}}{\sqrt{2\pi}\sigma_L P_R}, \tag{45.85}$$

where P_R is the received optical power and the variance σ_L^2 is calculated from the turbulence intensity. The average received optical power without considering turbulence is $P_{L0} = P_l P(\theta_s) A_R \cos(\xi) \exp(k_e L)/(4\pi L^2)$. Thus, the P_R edge of the PDF is

$$p(P_R) = \int p(P_l, P_R) dP_l = \int p(P_R|P_l) p(P_l) dP_l. \tag{45.86}$$

Ultraviolet Indirect Vision PDF Experimental Measurements

The probability density function was measured in sunny, rainy, and hazy weather conditions, in which the distance between the receiving and transmitting ends was 4 m and the transmission power was 20 mW. Figure 45.33 shows the probability density function of the measured ultraviolet light intensity under different transmission and reception elevation angle conditions.

As shown in Fig. 45.33, the probability density function of the received optical power is basically in line with the Gaussian model for different weather conditions. After normalizing the received power value, the variance of the probability density function differs under different weather conditions and transmission and reception elevation angles; however, the change trend is basically the same.

After the signal emitted by the transmitting end passes through the atmospheric channel, different transmission and reception elevation angles have different effects on the attenuation, and the magnitude of the path loss also differs. When the elevation angle of the signal is increased from $\beta_R = \beta_T = 20^\circ$ to $\beta_R = \beta_T = 40^\circ$, the energy of the signal decreases gradually, and the variance of the probability density function of the signal strength changes little. However, when the elevation angle of the signal increases to $\beta_R = \beta_T = 60^\circ$, the attenuation of the signal energy is more serious, that is, with the synchronous increase of β_R and β_T, the variance of the signal strength probability density function increases, and the energy attenuation gradually increases.

To reflect the distribution of light intensity under different receiving angles, the light intensity of different receiving elevation angles was measured, and the results are shown in Fig. 45.34. The probability density function of the ultraviolet light intensity was measured by changing the elevation angle of the receiving end. The distance between the receiving and transmitting ends was 4 m, the transmission power was 20 mW, and the elevation angle of the transmitting end was fixed at $\beta_T = 80^\circ$.

It can be seen from Fig. 45.34 that the probability density function of the received optical power conforms to the Gaussian model, and after normalizing the received power value, the variance of the probability density function differs under different transmission and reception elevation angles. After the signal emitted by the transmitting end passes through the atmospheric channel, different transmission and reception elevation angles have different effects on its attenuation, and the size of the path loss is also different.

With the increase of the receiving elevation angle β_R, the signal energy attenuation is very small, and the variance of the signal strength probability density function changes very little, indicating that the influence of the β_T/β_R ratio on the energy attenuation of the communication signal is greater. Therefore, when designing an ultraviolet optical transceiver system, the receiving elevation angle should be fixed as much as possible, and the emission elevation angle should be reduced to reduce the path loss.

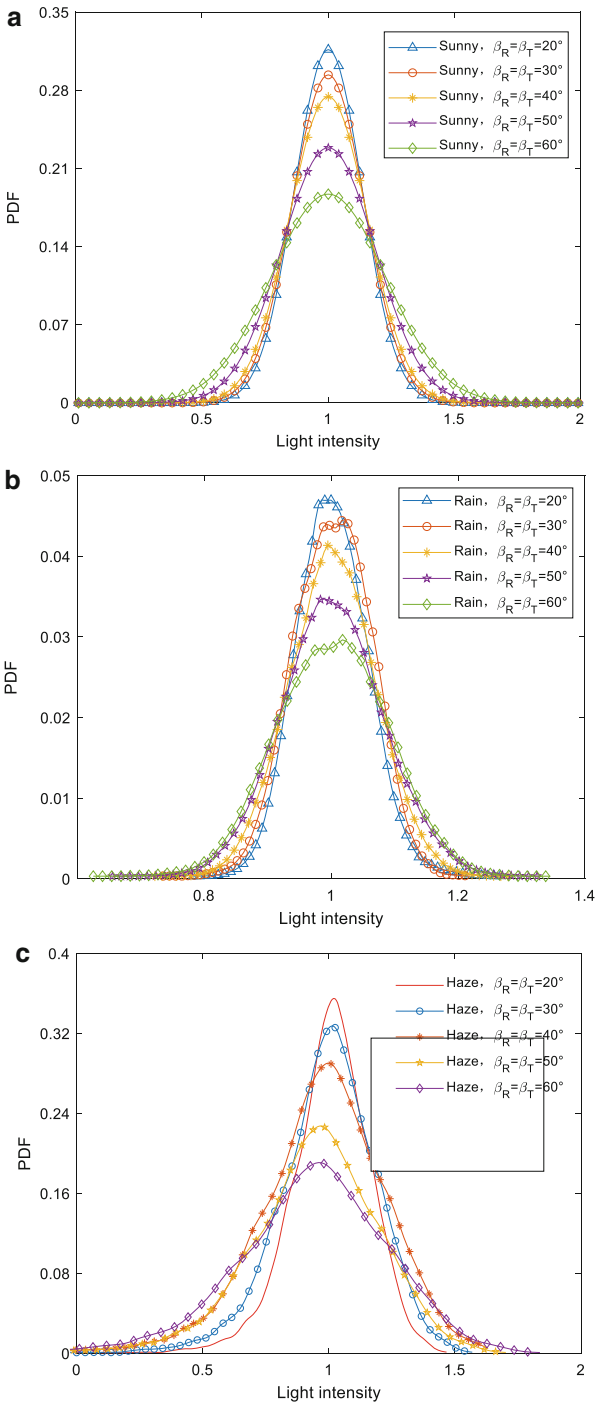


Fig. 45.33 Probability density function of light intensity at different elevation angles for different weather conditions. (**a**) Probability density function of light intensity on sunny days. (**b**) Light intensity probability density function on rainy days. (**c**) Light intensity probability density function on hazy days

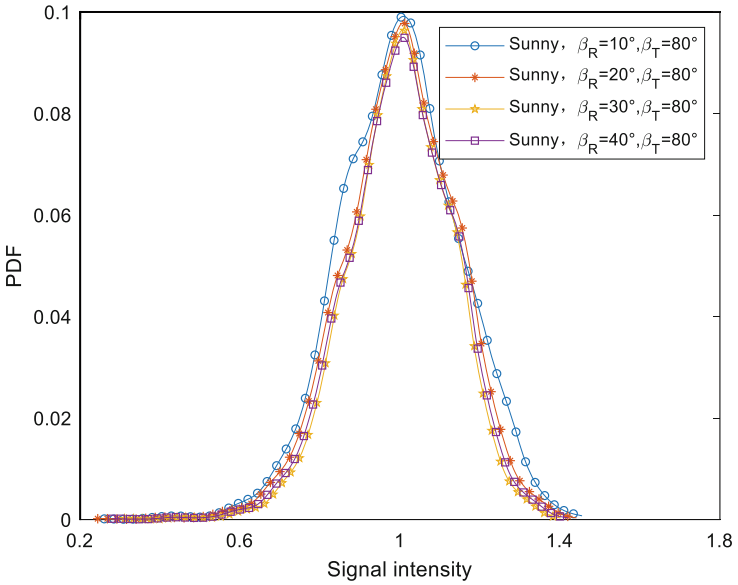


Fig. 45.34 Light intensity probability density function at different receiving elevation angles

Summary

In this study, an experimental measurement method was deduced, changes in ultraviolet background radiation at different times and seasons under sunny weather conditions were derived, and the effectiveness of the theoretical model was verified by comparison with the measured data. In addition, this study compared and analyzed the differences between the measured data and theoretical models in cloudy and hazy weather conditions.

Based on a theoretical model at the surface, the ultraviolet light noise distribution at high altitudes was measured, and the change in high-altitude noise was restored. The physical characteristics of aerosol particles in low-visibility environments, such as rainy days, foggy days, and dusty weather, were analyzed separately.

Moreover, a new empirical attenuation model was used to simplify the asymmetric factors in the scattering phase function for the common empirical model of optical attenuation for the analysis of ultraviolet atmospheric channel attenuation, complex calculation of the multiscattering transmission model, and short ultraviolet communication distance in low-visibility environments. Finally, the light intensity distribution characteristics of ultraviolet optical communication under different weather conditions and transmission and reception elevation angles were analyzed.

References

1. Ding Y, Fan J T, Tong S F, et al. Range extension method for communication distance of atmospheric channel ultraviolet optical communication system [J]. Acta Photonica Sinica, 2012, 41(9): 1047–1052.
2. Sunstein D E. A scatter communications link at ultraviolet frequencies [D]. Cambridge: Massachusetts Institute of Technology, 1968: 35–45.
3. Reilly D M. Atmospheric optical communications in the middle ultraviolet [D]. Cambridge: Massachusetts Institute of Technology, 1976:77–90.
4. Geller M, Keenan T E, Altman D E, et al. Optical non-line-of-sight covert, secure high data communication system [P]. U.S. Patent: 4,493,114. 1985:1–8.
5. Luettgen M R, Shapiro J H, Reilly D M. Non-line-of-sight single-scatter propagation model [J]. JOSA A, 1991, 8(12): 1964–1972.
6. Yuan R, Ma J. Review of ultraviolet non-line-of-sight communication [J]. China Communications, 2016, 13(6): 63–75.
7. Z. Xu and B. M. Sadler. Ultraviolet communications: potential and state- of-the- art [J]. Proceedings of IEEE, 2008, 5(46): 67–68.
8. Shaw G A, Siegel A M, Model J, et al. Recent progress in short-range ultraviolet communication [J]. Unattended Ground Sensor Technologies and Applications VII, 2005, 1(5796): 214–225.
9. Chen G, Abou-Galala F, Xu Z, et al. Experimental evaluation of LED-based solar blind NLOS communication links [J]. Optics express, 2008, 16(19): 15059–15068.
10. Ding H, Chen G, Majumdar A K, et al. Modeling of non-line-of-sight ultraviolet scattering channels for communication [J]. IEEE journal on selected areas in communications, 2009, 27(9): 1535–1544.
11. Li Y, Wang L, Xu Z, et al. Neighbor discovery for ultraviolet ad hoc networks [J]. IEEE Journal on Selected Areas in Communications, 2011, 29(10): 2002–2011.
12. Wang P, Xu Z. Characteristics of ultraviolet scattering and turbulent channels [J]. Optics letters, 2013, 38(15): 2773–2775.
13. El-Shimy M A, Hranilovic S. Spatial-diversity imaging receivers for non-line-of-sight solar-blind UV communications [J]. Journal of Lightwave Technology, 2015, 33(11): 2246–2255.
14. Ardakani M H, Heidarpour A R, Uysal M. Performance analysis of MIMO NLOS UV communications over atmospheric turbulence channels [A]. 2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)[C]. Doha, Qatar: IEEE, 2016, 4673(1): 1–5.
15. Ardakani M H, Heidarpour A R, Uysal M. Performance analysis of relay-assisted NLOS ultraviolet communications over turbulence channels [J]. Journal of Optical Communications and Networking, 2017, 9(1): 109–118.
16. Refaai A, Abaza M, El-Mahallawy M S, et al. Performance analysis of multiple NLOS UV communication cooperative relays over turbulent channels [J]. Optics express, 2018, 26(16): 19972–19985.
17. Arya S, Chung Y H. Amplify-and-forward multihop non-line-of-sight ultraviolet communication in the gamma–gamma fading channel [J]. Journal of Optical Communications and Networking, 2019, 11(8): 422–436.
18. Hasan Hariq S, Odabasioglu N. Spatial diversity techniques for non-line-of-sight ultraviolet communication systems over atmospheric turbulence channels [J]. IET Optoelectronics, 2020, 14(5): 327–336.
19. Bhatia V, Jain S, Garg K, et al. Performance analysis of RKHS based detectors for nonlinear NLOS ultraviolet communications [J]. IEEE Transactions on Vehicular Technology, 2021, 70(4): 3625–3639.
20. Arslan C H, Dagefu F T, Moore T J, et al. Optimization of ultraviolet communication links based on finite difference stochastic approximation [J]. Optics Express, 2022, 30(20): 36283–36296.

21. Ni G Q, Zhong S D, Liu L X, et al. Research on free atmosphere ultraviolet optical communication [J]. Optical Technique, 2000, 26(4): 297–303.
22. Lan T, Ni G Q. Simulation study on atmospheric transmission characteristics of ultraviolet communication [J]. Transactions of Beijing Institute of Technology, 2003, 27(4): 419–423.
23. Chang S, Yang J, Yang J, et al. The experimental research of UV communication [A]. Wireless Communications and Networks [C]. China, Wuhan: SPIE, 2004, 5284(1): 344–348.
24. Lan Y Z, Xiao S L, Luo Y J, et al. Research on transmission and modulation technology of solar blind ultraviolet optical communication system [J]. Optoelectronics, 2007, 25(03): 206–211.
25. Zhao M, Xiao S L, Wang X, et al. Research on ultraviolet optical communication system based on LED [J]. Laser & Optoelectronics Progress, 2010, 47(4): 19–24.
26. Zou Y, Xiao S L, Li G H, et al. Research on ultraviolet wireless optical communication system based on LED [J]. Optoelectronic Technique, 2012, 32(3): 145–149.
27. Ke X Z, He H, Chen X. A New MAC Layer Avoidance Algorithm for Ultraviolet Ad Hoc Communication Networks [J]. Photoelectron. Lasers, 2010, 21(7): 1002–1006.
28. Ke X Z, Chen J N, Hou Z M. Directional Media Access Control Protocol for Ultraviolet Non-line-of-Sight Communication [J]. Photoelectron. Journal of Lasers, 2011, 22(8): 1190–1195.
29. Zhao T F, Ke X Z. Monte Carlo Method for Modeling Non-Direct UV Light Scattering Coverage [J]. Acta Physica Sinica, 2012, 61(11): 285–296.
30. Wu C Y, Zuo Y, Fan C, et al. Research on MAC layer power control in ultraviolet optical communication network [J]. Journal of Optical Communication Research, 2013, 39(6): 28–31.
31. Xu C, Zhang H. Monte-Carlo based modeling for ultraviolet non-line-of-sight communication channels with typical obstacles [A]. 2015 International Conference on Wireless Communications & Signal Processing (WCSP)[C]. China Nanjing: IEEE, 2015, 4673: 1–5.
32. Yang G, Li X Y, Chen M, et al. A New Time-Division Multiple Access Neighbor Discovery Algorithm for Ultraviolet Ad Hoc Networks [J]. Photoelectron. Journal of Lasers, 2015, 26(6): 1074–1080.
33. Li J B, Wu X J, Wang H C, et al. Anti-jamming Relay Link Method and Power Demand Analysis of Ultraviolet Non-direct Communication [J]. Progress in Laser & Optoelectronics, 2015, 52(3): 99–105.
34. Song P, Song F, Li Y H, et al. Multi-user interference problem of non-direct UV optical communication networking [J]. Acta Photonica Sinica, 2016, 1(9): 159–166.
35. Song P, Zhou X L. Design and Communication Performance Analysis of Ultraviolet Mobile Ad Hoc Network Nodes [J]. Acta Optica Sinica, 2018, 38(3): 290–297.
36. Li J L, Zhao H, Qu G, et al. Performance Analysis of Different Wavelengths of Ultraviolet Optical Signals in High-altitude Atmospheric Communication [J]. Semiconductor Optoelectronics, 2017, 38(1): 79–84.
37. Wang P H, Wang J Y, Li J H, et al. Experiment on UV Light Communication Multidecimal Modulation System Based on Low Voltage Mercury Lamp [J]. Optical Communication Technology, 2019, 43(8): 21–24.
38. Zhang J, Zhou M Y. Design of a low-latency ultraviolet optical communication system [J]. Journal of Lasers, 2020, 41(4): 150–154.
39. Cao T, Gao X, Wu T, et al. Single-scatter path loss model of LED-based non-line-of-sight ultraviolet communications [J]. Optics letters, 2021, 46(16): 4013–4016.
40. Li C, Li J, Xu Z, et al. Research on the lossless contention MAC protocol and the performance of an ultraviolet communication network [J]. Optics Express, 2021, 29(20): 31952–31962.
41. Cao T, Wu T, Pan C, et al. Single-collision-induced path loss model of reflection-assisted non-line-of-sight ultraviolet communications [J]. Optics Express, 2022, 30(9): 15227–15237.
42. Chen G, Xu Z, Sadler B M. Experimental demonstration of ultraviolet pulse broadening in short-range non-line-of-sight communication channels [J]. Optics express, 2010, 18(10): 10500–10509.

43. Z Yong, L Yanfeng, W Dengjia, etc. Applicability analysis of daily total solar radiation calculation model in different climatic regions of China and optimization of General Calculation Model [J]. Journal of solar energy, 2022, 43(09): 1–7
44. Kasten F, Czeplak G. Solar and terrestrial radiation dependent on the amount and type of cloud [J]. Solar energy, 1980, 24(2): 177–189.
45. Prescott J A. Evaporation from a water surface in relation to solar radiation [J]. Trans. Roy. Soc. S. Aust., 1940, 46(1): 114–118.
46. Chen S, Hinson J W, Lee J, et al. Branching fraction and photon energy spectrum for b→ s γ [J]. Physical Review Letters, 2001, 87(25): 13–15.
47. Fu W B. Multi-photon process [J]. Advances in laser and optoelectronics, 1984, 21(1): 15.
48. Huang Z X. Some problems in the theory and application of single photon technology [J]. Journal of Communication University of China (natural science edition), 26(02), 2019: 1–18.
49. Einstein A. Concerning an heuristic point of view toward the emission and transformation of light [J]. American Journal of Physics, 1965, 33(5): 367.
50. Kirchhoff G. I. On the relation between the radiating and absorbing powers of different bodies for light and heat [J]. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1860, 20(130): 1–21.
51. Reda I, Andreas A. Solar position algorithm for solar radiation applications [J]. Solar energy, 2004, 76(5): 577–589.
52. Grena R. An algorithm for the computation of the solar position [J]. Solar Energy, 2008, 82(5): 462–470.
53. Emilio M, Kuhn J R, Bush R I, et al. Measuring the solar radius from space during the 2003 and 2006 Mercury transits [J]. The Astrophysical Journal, 2012, 750(2): 135.
54. Ackermann A S E. The Utilisation of Solar Energy [J]. Journal of the Royal Society of Arts, 1915, 63(3258): 538–565.
55. Wong L T, Chow W K. Solar radiation model [J]. Applied energy, 2001, 69(3): 191–224.
56. Fang X J. Temporal variation of atmospheric transparency coefficient in our country [J]. Journal of Geography, 1988, 43(01): 52–59.
57. Hou Y L, Huang Y F. Analysis of the causes of changes in ultraviolet radiation [J]. Journal of Heze Normal College,2004, 26(04): 70–71.
58. Liu J M, Ding Y G, Huang Y D, et al. Correlation analysis of solar ultraviolet radiation intensity and meteorological elements [J]. Plateau Meteorology, 2003, 22(01): 45–50.
59. Wang X, Wen J, Liu R, et al. Effects of precipitation process on near-ground solar radiation in Badain Jaran Desert [J]. Journal of Arid Meteorology, 2011, 29(4): 427–432.
60. Hu L Q, Ji H Y. Research on regional atmospheric transparency in Xinjiang [J]. Journal of Desert Research, 2008, 28(2): 332–337.
61. Planck M. On the law of distribution of energy in the normal spectrum [J]. Annalen der physik, 1901, 4(553): 1–2.
62. Coskun C, Oktay Z, Dincer I. Estimation of monthly solar radiation distribution for solar energy system analysis [J]. Energy, 2011, 36(2): 1319–1323.
63. Xu Z, Sadler B M. Ultraviolet communications: potential and state-of-the-art [J]. IEEE Communications Magazine, 2008, 46(5): 67–73.
64. Razavi M, Shapiro J H. Optical wireless communications via diversity reception and optical preamplification [J]. IEEE Transactions on wireless Communications, 2005, 4(3): 975–983.
65. Luettgen M R, Shapiro J H, Reilly D M. Non-line-of-sight single-scatter propagation model [J]. JOSA A, 1991, 8(12): 1964–1972.
66. Reilly D M, Warde C. Temporal characteristics of single-scatter radiation [J]. JOSA, 1979, 69(3): 464–470.
67. Ke X Z. Theory of ultraviolet self-organizing network [M]. Beijing: Science Press, 2011: 40–42.
68. Song P, Liu C, Zhu L, et al. Single scattering path loss analysis of ultraviolet optical wireless communication in mobile scenarios [J]. Acta Optica Sinica, 2020, 40(04): 28–37.

69. Hale G M, Querry M R. Optical constants of water in the 200-nm to 200-μm wavelength region [J]. Applied optics, 1973, 12(3): 555–563.
70. Liu L, Li H, Gao T C. Approximate ellipsoidal model of raindrops and its near-infrared scattering characteristics [J]. Journal of Meteorological Sciences, 2008, 28(3): 271–275.
71. Hou Q, L X Yi, Hou Z H, et al. Influence of Different Weather on Non-line-of-Sight Transmission of Ultraviolet Light [J]. Communication Confrontation, 2011(3): 13–15.
72. Song Z F. Fundamentals of Applied Atmospheric Optics [M]. Beijing: China Meteorological Press, 1990: 17–20.
73. Lin Y, Xu Z Y, Wang J Y, et al. Study on scattering transmission characteristics of non-line-of-sight atmosphere in fog environment [J]. Acta Optica Sinica, 2013, 33(9): 11–14
74. Liu Bo, Li P C, Li J, et al. Bit error rate performance of free-space optical communication system under different fog concentrations [J]. Electronic Information Countermeasures Technology, 2019, 34(6): 44–47.
75. Zhong H, Zhou J, Du Z, et al. A laboratory experimental study on laser attenuations by dust/sand storms [J]. Journal of Aerosol Science, 2018, 121: 31–37.
76. Wang H X, Zhu Y Z, Tian T, et al. Study on transmission characteristics of laser in different types of aerosols [J]. Acta Physica Sinica, 2013, 62(2): 324–333

Research Progress on OWC/RF Hybrid Communication Systems

46

Contents

Introduction 1704
Foreign and Domestic Research Progress 1704
Foreign Research Progress 1705
Domestic Research Progress 1707
System Model 1709
OWC/RF Hybrid Parallel System Model 1709
OWC/RF Hybrid Relay System Model 1712
Multiuser Hybrid Relay System 1713
Intelligent Reflecting Surface-Assisted Hybrid Relay System 1713
Application Scenarios for OWC/RF Hybrid Communication 1714
Hybrid System for Space Communication 1714
Hybrid Systems in Vehicle Communications 1715
Hybrid Systems in Underwater Communication 1716
Hybrid Systems in 5G Retransmission 1718
Hybrid System Performance 1719
OWC Subsystem 1719
RF Subsystem 1720
Hybrid System 1720
Problems, Challenges, and Solutions 1721
Weather Conditions 1721
Atmospheric Turbulence 1722
Pointing Errors 1722
Multipath Effect 1723
Summary 1723
References 1724

Abstract

The optical wireless/radio frequency hybrid communication system makes up for the limitations of a single communication system, makes up for the shortcomings of radio frequency communication while maximizing the advantages of optical wireless communication, significantly improves the availability and reliability of the link, and provides a unique solution for future high-speed and high-capacity

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_46

information transmission. This chapter reviews the research progress of optical wireless/radio frequency hybrid communication system in recent years, firstly introduces the working principles and applications of different hybrid communication system models, then derives the outage probability and average bit error rate of the hybrid system, then discusses the challenges and solutions of the hybrid communication system, and finally looks into the future development trend of optical wireless/radio frequency hybrid communication system, aiming to lay the foundation for further research and development of optical wireless/radio frequency hybrid communication system.

Keywords

Optical wireless/radio frequency hybrid communication · Atmospheric turbulence · Switching mechanism · Relay protocol

Introduction

Hybrid optical wireless communication/radio frequency (OWC/RF) systems make up for the limitations of a single communication system. They maximize the advantages of optical wireless communication while making up for the shortcomings of radio frequency communication. They significantly improve link availability and reliability and provide a unique solution for future high-speed and high-capacity information transmission.

Optical wireless communication (OWC) [1] uses light as a carrier wave for transmissions in space. It has high speed, high capacity, and high confidentiality, does not need for spectrum licensing, and has received considerable attention in recent years. However, OWC is affected by atmospheric turbulence, weather conditions (fog, snow, etc.), and pointing errors but is not sensitive to rain and multipath fading [2].

Radio frequency (RF) communication is mainly affected by rain and multipath fading but is insensitive to turbulence, pointing errors, snow, and fog [3]. It is precisely complementary to OWC, and OWC/RF hybrid communication has emerged. The OWC/RF hybrid communication system combines the advantages of OWC and RF communication and provides higher reliability and availability than a single communication system under highly variable channel conditions. It meets the requirements for future information transmissions with high capacity, high speed, long distance, safety, and reliability. It is an important development direction for optical wireless communication.

Foreign and Domestic Research Progress

The OWC/RF hybrid communication system is a new scheme to address the impact of severe weather on wireless communication systems. Because of the effects of different weather on OWC and RF communication, they are combined to form an

OWC/RF hybrid communication system that improves the system reliability and realizes high-quality end-to-end communication.

Foreign Research Progress

In 2009, Schober et al. [4] proposed a new OWC/RF hybrid system architecture to enable the acquisition of channel state information. Bit-interleaved coded modulation (BICM) was used on the bit stream transmitted by the RF and OWC subchannels to improve the robustness of the hybrid system.

In 2009, Nadeem et al. [5] considered the high dependence of OWC links on weather conditions and formed a hybrid network using GHz RF links as backup for OWC links. They investigated the effect of fog, rain, and snow on the hybrid OWC/GHz RF network to select the GHz-RF links with the best complementary characteristics. The simulation results showed that a 40-GHz RF link significantly improved the availability of the hybrid system link.

In 2010, Abdulhussein et al. [6] proposed an automatic repeat request (ARQ) scheme using rate-free coding in an OWC/RF hybrid system. Regardless of changes in the OWC or RF channel conditions, this scheme fully utilized the available OWC and RF channel resources without requiring a priori rate selection at the transmitter. The simulation results showed that using rate-free coding was more advantageous than using fixed rate coding under strong atmospheric turbulence conditions.

In 2011, Lee et al. [7] investigated the performance of an asymmetric dual-hop relay OWC/RF hybrid system, in which the RF channel followed Rayleigh fading and the OWC channel followed gamma–gamma fading. The relay node converted the received RF signal into an optical signal using subcarrier intensity modulation. Simulation results showed that the interrupt performance of this hybrid OWC/RF link was extremely close to that of an RF/RF link as the signal-to-noise ratio increased.

In 2013, Ansari et al. [8] analyzed a hybrid OWC/RF communication system with pointing errors and derived the cumulative distribution, probability density, and moment-generating functions for the system, which led to an expression for the average bit error rate (BER) of the hybrid system. Monte Carlo simulation results showed that the system BER performance gradually deteriorated as the pointing errors increased.

In 2014, Usman et al. [9] studied a hybrid OWC/RF communication system based on hard-switching technology and compared the performance of the hybrid communication system under both single and dual OWC threshold conditions. Numerical simulations showed that using OWC/RF hybrid communication effectively improved the system performance, and the dual OWC threshold switching scheme achieved performance similar to a single OWC threshold. It avoided frequent link switching and extended the OWC link lifetime.

In 2016, Touati et al. [10] investigated the effect of joint fading on hybrid OWC/RF links. The experimental results showed that the hybrid OWC/RF system

performed significantly better than OWC alone, in the cases of strong turbulence or a low signal-to-noise ratio, which effectively overcame the drawbacks of OWC links.

In 2017, Hasna et al. [11] proposed an adaptive and efficient OWC/RF hybrid communication system considering the actual communication environment. It could multiplex multiple RF users into a single high-speed OWC link and thus achieved a higher system capacity. However, the application of multiuser diversity only slightly improved the system performance under low signal-to-noise ratio conditions.

In 2019, Amirabadi et al. [12] proposed a novel OWC/RF hybrid system model using a receive-diversity technique in a parallel single-hop OWC/RF hybrid system. Simulation results showed that the proposed OWC/RF hybrid system had a lower outage probability than normal OWC and RF systems with receive diversity under different average signal-to-noise ratio conditions. However, its reduction in power consumption came at the cost of increased receiver complexity.

In 2019, Esmail et al. [13] experimentally evaluated all-optical OWC/RF hybrid communication links for 5G applications under harsh sandy and dusty conditions. The experimental results showed that the RF link had good communication performance in a dense sandy environment, while the OWC link experienced communication interruptions, proving that an OWC/RF parallel link was a feasible solution to overcome the dust effect.

In 2020, Ashrafzadeh et al. [14] constructed a new hybrid three-hop-relay RF-OWC system to solve the problems of congestion and nonline-of-sight transmission in urban environments and mitigate the fading effects caused by turbulence. However, the system relay node had to be equipped with partial or complete OWC transceivers, significantly increasing the system complexity.

In 2021, Khan et al. [15] analyzed an adaptive OWC/RF hybrid system under different weather conditions. Using adaptive modulation and power allocation, the adaptive OWC/RF hybrid system achieved more than 1 dB performance gain at a minimum power level and adapted to changing weather conditions.

In 2022, Vishwakarma et al. [16] investigated the ergodic capacity performance of a hybrid OWC/RF system based on an adaptive combining switching scheme. Closed expressions for the ergodic capacity were derived using a generalized binary Meijer G-function and binary Fox H-function under high signal-to-noise ratio conditions. The optimal switching threshold signal-to-noise ratio and beamwidth values were determined to obtain the optimal ergodic capacity of the adaptive merged system. The numerical results showed that the ergodic capacity of the adaptive combining-based OWC/RF hybrid system was greater than that of the maximal ratio combining (MRC) and hard-switching-based OWC/RF hybrid communication system.

In 2022, Magidi et al. [17] analyzed the performance of a hybrid OWC/RF communication system under the combined effects of atmospheric fading and pointing errors. In this case, a 16-way quadrature amplitude modulation (16QAM) was used for the RF link, and multipulse position modulation (MPPM) was used for the OWC link. The experimental results showed that the smaller the normalized jitter standard deviation, the better the outage probability performance, regardless of the atmospheric turbulence severity. Using an OWC/RF hybrid link reduced the outage probability from the 10^{-2} order of magnitude for the OWC link alone to the 10^{-5} order of magnitude.

Domestic Research Progress

In 2008, Shu et al. [18] studied the theory of automatic link switching for OWC/RF hybrid communication systems. The system signal-to-noise ratio was first improved by reducing the communication rate when the weather conditions deteriorated; and then the system switched to RF channels if the performance of the low-speed OWC communication system was still below a set threshold. This method made full use of the complementary advantages of OWC and RF links to reduce the amount of channel switching while improving the system reliability.

In 2009, Meng [19] demonstrated that the received power strength could reflect the channel quality. They improved the rapid spanning tree protocol (RSTP) used in switches to handle redundant network links, as a switching method for OWC/RF hybrid systems, effectively improving the system communication quality.

In 2010, Su et al. [20] studied the switching technique of an OWC/RF hybrid communication system based on the network packet loss rate. The switching judgment mechanism more accurately responded to the current link conditions, so that the hybrid link was switched in a timely and effective manner to ensure that the data could be transmitted without interruption.

In 2015, to alleviate the influence of multipath fading and atmospheric turbulence on an OWC/RF hybrid system, Gao et al. [21] adopted diversity transmission at the source, selective combination at the sink and only one RF receiving antenna and relay aperture for information forwarding at the relay node. This provided a theoretical reference for the design of an OWC/RF hybrid communication system.

In 2016, Wang et al. [22] extended multiuser diversity to OWC/RF hybrid links and proposed a point-to-multipoint system consisting of a hybrid access point, multiple OWC users, and multiple RF users. For each transmission module, the user with the best link quality was selected to communicate with the hybrid access point through a user scheduling policy.

In 2017, Zhao et al. [23] established an OWC/RF hybrid communication system model with aperture smoothing using subcarrier multiband phase-shift keying modulation. They compared and analyzed the outage probability and bit error rate performance of the hybrid system under different turbulence intensities. The simulation results showed that the aperture-smoothing effect significantly improved the communication performance of the hybrid system and the OWC channel played a major role in the hybrid system performance.

In the same year, Zhao et al. [24] studied hybrid systems with RF and OWC links subject to the Nakagami-m and exponential Weibull distributions, respectively. Based on end-to-end statistical characteristics, a new outage probability expression was derived to evaluate the performance of fixed-gain relay systems and channel state information-assisted relay systems. It was found that the latter was very sensitive to atmospheric turbulence.

In 2017, Shao et al. [25] investigated the system performance, using low-density parity-check (LDPC) hybrid coding and pulse-position modulation for the OWC link and quadrature amplitude modulation for the RF link. They used a soft-switching

mechanism for the reliable transmission of physical layer data in OWC/RF hybrid communication link networking. Simulation results showed that the performance of the proposed soft-switching hybrid coding-modulation scheme improved from 1.3 dB to 8.0 dB, which significantly improved the data transmission reliability of the OWC/RF hybrid parallel system.

In 2019, Yi et al. [26] studied the performance of a decode-and-forward-based dual-hop OWC/RF hybrid communication system in a 5G cloud radio access network (C-RAN). The user transmitted data to a remote RF unit via an RF link, and the RF unit acted as a relay system to forward the received data to the baseband unit. Owing to the presence of small-scale fading and line-of-sight components, the RF link obeyed a κ-μ shadow distribution.

In 2020, Liu et al. [27] introduced a dual-hop OWC/RF hybrid system into deep space communication. The signal arrived at the relay system through the RF link, was amplified and forwarded by the relay system, and then was transmitted by the OWC link to reach the receiver. Compared with traditional OWC systems, this OWC/RF hybrid system can achieve a lower BER in deep space communication. The system performance can be further improved when the hybrid system adopts hybrid modulation.

In 2020, Liu et al. [28] studied the performance of an RF/OWC hybrid aviation communication system based on selective combined diversity technology. The system transmitter did not require feedback links to obtain the channel state information, simplifying system design. Under a high signal-to-noise ratio, the RF/OWC hybrid parallel communication scheme with reception diversity had better performance than that with low-order modulation.

In 2021, Wu et al. [29] investigated the advantages of subcarrier modulation technology in atmospheric turbulence channels. The BER performance of the following OWC/RF hybrid system subcarrier modulation techniques was investigated: differential binary phase-shift keying (DBPSK), noncoherent binary frequency-shift keying (NBFSK), coherent binary phase-shift keying (CBPSK), and coherent binary frequency-shift keying (CBFSK). The simulation results showed that the performance of the OWC/RF hybrid communication system using CBPSK subcarrier modulation technology was obviously better than that of the other three modulation technologies.

In 2022, Liu et al. [30] proposed a millimeter-wave RF/multiple-input multiple-output (MIMO) OWC relay system based on the RF/OWC hybrid relay system. It used a fixed-gain scheme at the relay and an equal-gain merging method at the receiver to combine multiple signals. Numerical simulations demonstrated that the MIMO technique effectively mitigated the performance degradation of the hybrid relay system due to the enhanced atmospheric turbulence intensity. The performance of the hybrid relay system was gradually enhanced as the number of transceiver apertures of the OWC link increased; however, the system complexity also increased. Considering the system complexity and cost, the hybrid system performance was optimized with two transmitting and two receiving apertures.

System Model

OWC/RF Hybrid Parallel System Model

OWC/RF hybrid parallel systems are critical to improving the reliability of hybrid systems by selecting the appropriate channel according to the current communication conditions. Channel selection requires an efficient link-switching mechanism [31], which currently includes hard- and soft-switching types.

Hard Switching

An OWC/RF hybrid parallel system based on a hard-switching mechanism [32] means that either the OWC link or the RF link can be selected for data transmission at a given moment. Typically, the OWC link is used as the primary link, and the RF link is used as the backup link. The working principle is shown in Fig. 46.1.

The OWC link is used by default when the system starts up, and the link is automatically switched for information transmission, according to the threshold information detected by the receiving end of the OWC link. The switching threshold of the hard-switching mechanism can be set according to the system signal-to-noise ratio, bit error rate, and power [32–35]. Most current studies are based on the signal-to-noise ratio threshold. Taking the signal-to-noise ratio threshold as an example, the working principles of single and dual-threshold handovers for OWC links in hybrid systems are described below.

Single Threshold Switching

The signal-to-noise ratio thresholds for the OWC link and RF link are set to $\gamma_{\text{th}}^{\text{OWC}}$ and $\gamma_{\text{th}}^{\text{RF}}$, respectively, and the operating region of the single-threshold hybrid system is shown in Fig. 46.2. If the instantaneous signal-to-noise ratio of the OWC link is higher than the threshold $\gamma_{\text{th}}^{\text{OWC}}$ (in region 1), the OWC link is used to transmit information. If the instantaneous signal-to-noise ratio of the OWC link is lower than $\gamma_{\text{th}}^{\text{OWC}}$, the system will check the communication quality of the RF link. If the instantaneous signal-to-noise ratio of the RF link is higher than the threshold $\gamma_{\text{th}}^{\text{RF}}$ (within region 2), it switches to the RF link to transmit the information.

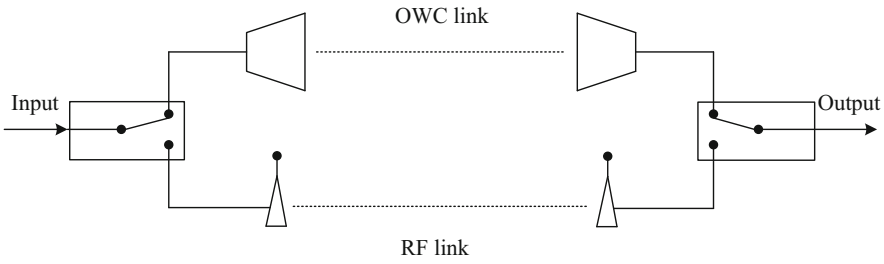


Fig. 46.1 Hard-switching schematic [16]

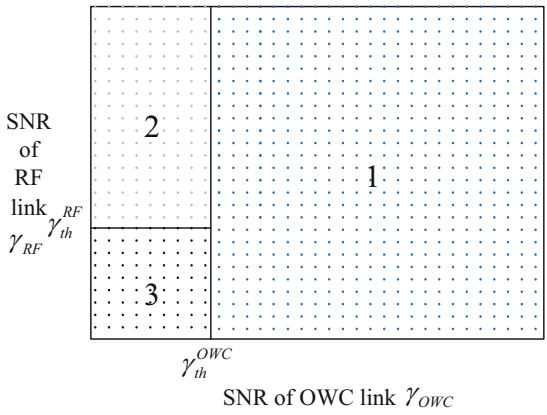


Fig. 46.2 Working area of a hybrid system under single-threshold hard switching [9]

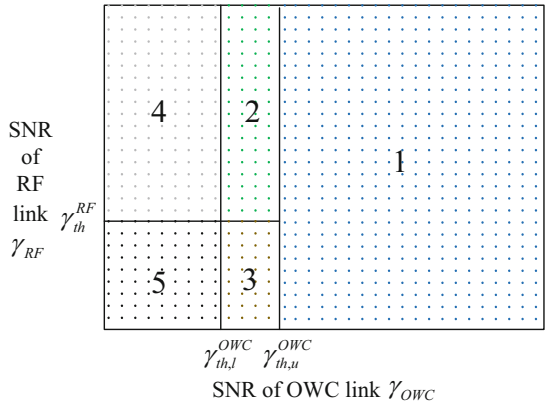


Fig. 46.3 Working area of the hybrid system under dual-threshold hard switching [9]

If the instantaneous signal-to-noise ratios of both links are below both thresholds (in region 3), the communication is interrupted. The single-threshold switching scheme has the problem of frequent link switching. To reduce the frequency of link switching and extend the link life, the double-threshold switching scheme can be used.

Dual-Threshold Switching

Dual-threshold switching refers to setting two signal-to-noise ratio thresholds on the OWC link, namely, the upper limit $\gamma_{\mathrm{th},u}^{\mathrm{OWC}}$ and the lower limit $\gamma_{\mathrm{th},l}^{\mathrm{OWC}}$ ($\gamma_{\mathrm{th},l}^{\mathrm{OWC}} < \gamma_{\mathrm{th},u}^{\mathrm{OWC}}$). Another signal-to-noise ratio threshold $\gamma_{\mathrm{th}}^{\mathrm{RF}}$ is set on the RF link, dividing the plane formed by the signal-to-noise ratio values of the OWC/RF hybrid system into five regions, as shown in Fig. 46.3.

If the instantaneous signal-to-noise ratio of the OWC link is higher than $\gamma_{\text{th},u}^{\text{OWC}}$ (in region 1), the OWC link is used to transmit the information. If the instantaneous signal-to-noise ratio of the OWC link is between $\gamma_{\text{th},l}^{\text{OWC}}$ and $\gamma_{\text{th},u}^{\text{OWC}}$, and the instantaneous signal-to-noise ratio of the RF link is higher than $\gamma_{\text{th}}^{\text{RF}}$ (in region 2), either the RF link or the OWC link may be used, depending on the previous state of the OWC link. If the instantaneous signal-to-noise ratio of the OWC link is higher than $\gamma_{\text{th},u}^{\text{OWC}}$ at the previous moment, the information is transmitted by the OWC link and vice versa.

If the instantaneous signal-to-noise ratio of the OWC link is between $\gamma_{\text{th},l}^{\text{OWC}}$ and $\gamma_{\text{th},u}^{\text{OWC}}$, and the instantaneous signal-to-noise ratio of the RF link is lower than $\gamma_{\text{th}}^{\text{RF}}$ (in region 3), the probability of the OWC link switching is the same as in region 2; however, the RF link always remains off in this region, so either the information is transmitted using the OWC link in region 3, or the communication is interrupted.

If the instantaneous signal-to-noise ratio of the OWC link is lower than $\gamma_{\text{th},l}^{\text{OWC}}$ and the instantaneous signal-to-noise ratio of the RF link is higher than $\gamma_{\text{th}}^{\text{RF}}$ (in region 4), the RF link is used for transmission. If the instantaneous signal-to-noise ratios of the OWC link and the RF link are lower than $\gamma_{\text{th},l}^{\text{OWC}}$ and γ_{th}^{RF} (in region 5), respectively, both links are off and communication is interrupted.

The dual-threshold switching hybrid system sets two switching thresholds in the OWC link. The size of the gap between these two switching thresholds directly determines the frequency of link switching: the larger the gap between $\gamma_{\text{th},l}^{\text{OWC}}$ and $\gamma_{\text{th},u}^{\text{OWC}}$, the lower the link-switching frequency [9]. Hence, the gap between $\gamma_{\text{th},l}^{\text{OWC}}$ and $\gamma_{\text{th},u}^{\text{OWC}}$ should be increased as much as possible to ensure the performance of the hybrid system. This effectively reduces the amount of switching between the OWC and RF links and expands the life cycle of the link.

Soft Switching

The hard-switching-based OWC/RF hybrid parallel system allows only one channel to transmit at a time. The other channel is idle, leading to wasted resources and low bandwidth utilization. To overcome these drawbacks, a hybrid OWC/RF parallel system based on a soft-switching mechanism is proposed [36]. Its working principle is shown in Fig. 46.4.

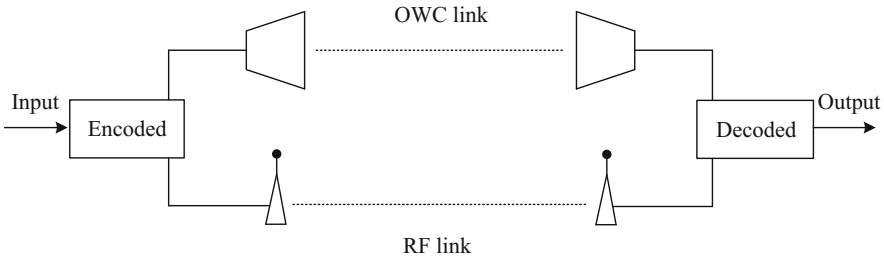


Fig. 46.4 Soft-switching schematic [16]

It uses a pair of coders and decoders to divide the data according to the real-time channel state information and assigns them to the two channels at different data rates. Hence, the data are transmitted on both OWC links and RF links [37], which improves the channel resource utilization and increases the channel capacity.

OWC/RF Hybrid Relay System Model

OWC/RF hybrid relay systems [38] provide wider coverage by relaying information from RF links to OWC links through relay nodes. Current research on hybrid OWC/RF relay systems mainly focuses on two-hop hybrid relay systems [23], where one hop uses an RF link transmission and the other hop uses an OWC link. The structure of a typical two-hop hybrid relay system is shown in Fig. 46.5.

At the source node S transmitter, the source data is modulated into an electrical signal, which is transmitted to the relay node R through an RF link. The relay node R modulates the received RF signal to an optical signal, and the optical signal is transmitted through the atmospheric channel to reach the destination node D.

Depending on the relaying protocols, they can be classified into amplify-and-forward (AF) and decode-and-forward (DF) relay systems [39, 40]. In the amplify-and-forward relay system, the relay node R directly amplifies the signal received from the source node S with a certain gain and forwards it to the destination node D. In the decode-and-forward relay system, the relay node R first decodes the signal received from the source node S and then recodes the processed signal and forwards it to the destination node D.

Because the decode-and-forward relay system requires the signal to be decoded and recoded, it is more costly and complex compared to the amplify-and-forward

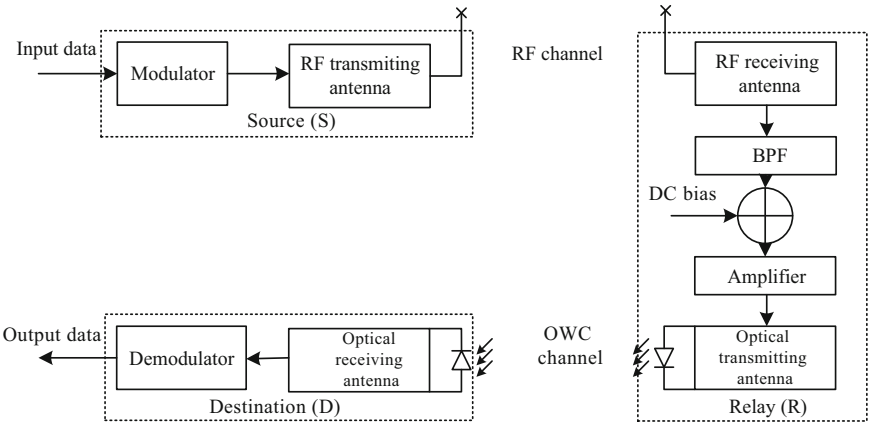


Fig. 46.5 Typical dual-hop OWC/RF hybrid relay system structure [41]

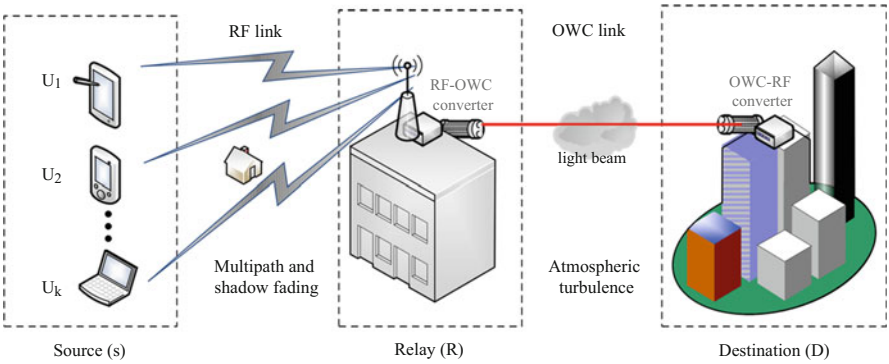


Fig. 46.6 Multiuser dual-hop hybrid relay system diagram [42]

relay system. However, the amplify-and-forward relay system amplifies the noise in the channel along with the signal, which leads to poor communication performance [41].

Multiuser Hybrid Relay System

The multiuser hybrid relay system [42] is shown schematically in Fig. 46.6. The source node S contains K mobile users, each equipped with a single antenna. There is no direct communication link between the user and the destination; the user can only send data to the destination through the relay node.

The model contains two types of communication links: a source-to-relay (S-R) RF link and a relay-to-destination (R-D) free-space optical link. In the RF link, the RF signal is affected by multipath and shadowing effects. In the free-space optical link, the optical signal is mainly affected by atmospheric turbulence and pointing errors. To provide a more accurate line-of-sight connection for OWC links, optical transmitters and receivers are generally placed on top of buildings.

Multiuser hybrid relay systems can simulate a number of practical applications [43], including the following;

1. Cellular communications, where mobile nodes send their data to a relay station, which forwards the data to a base station
2. Small cell networks, where mobile nodes on the floor of a building send their data to an access point, which forwards the information to a macro base station

Intelligent Reflecting Surface-Assisted Hybrid Relay System

An intelligent reflecting surface (IRS) [44] is a new type of device composed of a large number of low-cost passive components. By changing the amplitude or phase

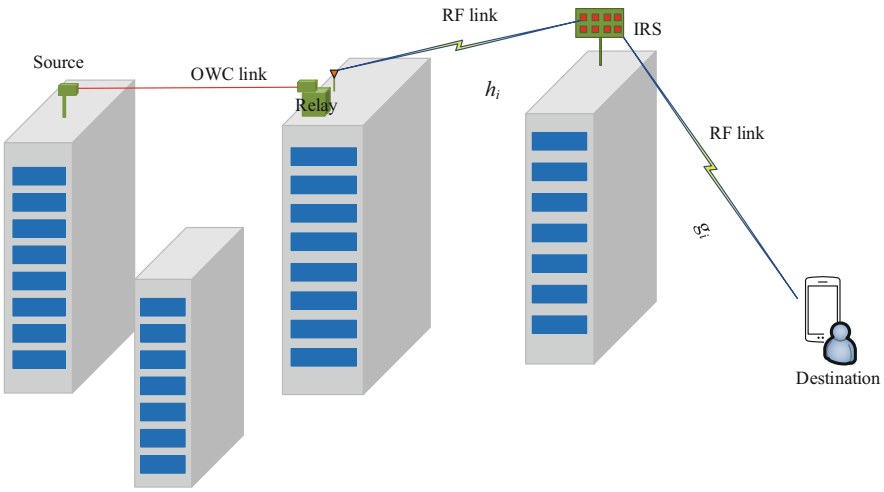


Fig. 46.7 Intelligent reflecting surface-assisted hybrid relay system diagram [45]

of the incident signal, it controls the phase distribution of the wavefront, allowing the incident beam to be reflected in any desired direction.

Owing to the small signal coverage of the intelligent reflecting surfaces, and the lack of a power supply and complex encoding and decoding algorithms, there is no mutual interference between the components of the intelligent reflecting surface. Thus, it can easily be integrated into wireless communication systems. The intelligent reflecting surface-assisted hybrid relay system model [45] is shown in Fig. 46.7.

An intelligent reflecting surface is installed on the top floor of a building. It utilizes RF links to reflect signals from the relay station to a predetermined receiver, solving the obstruction problem and ensuring reliable and high-speed data transmissions.

Application Scenarios for OWC/RF Hybrid Communication

Hybrid System for Space Communication

The US Air Force Research Laboratory conducted the Integrated RF/Optical Networked Tactical Targeting Networking Technologies (IRON-T2) program [46] to demonstrate the effectiveness of airborne hybrid OWC/RF communication. The project adapts terminals by adding high-speed routers and optical inputs, which allow interactive operations with supporting OWC link terminals. The router is designed to accommodate brief interruptions caused by flicker, and an RF channel provides retransmission signals for the high data-rate OWC systems, resulting in error-free transmissions.

The Deutsches Zentrum für Luft- und Raumfahrt (DLR) has developed a hybrid OWC/RF system in the VABENE project [47]. It combines an optical communication system developed by the DLR Institute for Communication and Navigation and a high-resolution camera sensor developed by the Institute for Remote Sensing. It is designed for monitoring and traffic management in large-scale events and natural disasters.

A simulated demonstration of a 1-Gbit/s downlink between a satellite camera sensor, the roof of the Institute of Communications and Navigation, and a service tower was conducted. The experimental results showed that small spacecraft could provide higher data rates through stable optical links, while RF links were independent of environmental conditions and compensated for the lack of optical communication links.

To achieve long-range and high data-rate communications, the National Aeronautics and Space Administration (NASA) Glenn Research Center designed the Teletenna, a hybrid telescope antenna [48], to support deep space communications. It combines optical and RF technology to alleviate engineering challenges and achieve 10–100 times higher data transmissions for deep space optical communications. The OWC link and RF link data rates are up to 84 Mbit/s and 287 Mbit/s, respectively. The Teletenna integrates a co-axial telescope and a Ka-band RF antenna with Cassegrain geometry [49], which is designed to improve the quality of communication systems in deep space missions. The system stability is improved by isolating the reflective and optical elements.

Hybrid Systems in Vehicle Communications

In recent years, research on network connectivity using low earth orbit (LEO) satellites has been increasing, and using OWC/RF hybrid networks for last-mile access in vehicular networking has received widespread attention [50]. Nguyen et al. [51] developed a transmission control protocol (TCP) performance analysis model, shown in Fig. 46.8. They considered an end-to-end network scenario in which a vehicle downloads data from a remote server (TCP source) via LEO satellite Internet access. The satellite is connected to a TCP source placed on an Internet cable.

The physical layer of the system consists of an OWC link and an RF link, both of which are channel coded using Reed–Solomon codes (RS) to improve the bit error performance of the physical layer. The link layer ensures reliable data transmission by selecting repeated automatic retransmission requests and resending the error frames.

In a hybrid OWC/RF vehicular self-assembly network, Gong et al. [52], from the University of Science and Technology of China, used distributed channel assignments and rate control to solve the cross-layer design problem and maximize network throughput. Owing to the limited channel resources in vehicle communication, to ensure the quality of service, the proposed channel allocation algorithm determines the channel usage through linearly programming the channel requirements of the measurement link, and the algorithm converges quickly. Then, it uses a

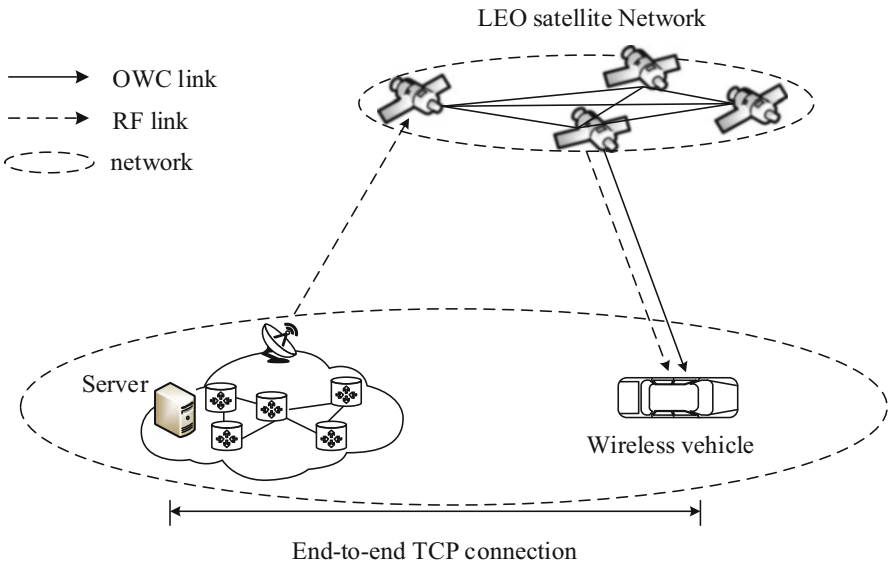


Fig. 46.8 Last-mile satellite Internet access scenario using vehicle Internets [50]

rate adjustment algorithm (the alternate direction multiplier method, ADMM) to control the transmission rate on each path.

Chowdhury et al. [53] demonstrated that OWC/RF hybrid communication can be used to improve the reliability of vehicle-to-vehicle (V2V) communication through the following three scenarios:

1. When a traffic accident occurs on foggy days, the front vehicle cannot transmit the accident information to the rear vehicle because the rear vehicle's optical receiver cannot detect the optical signal. In this case, the RF link can be used for communication.
2. OWC supports long-distance communication between vehicles. When the distance between two vehicles is longer than the RF communication range, they can switch to the OWC link.
3. When the vehicle in front turns left or right, a nonline-of-sight situation will occur, and the optical wireless communication link will be interrupted. At this time, RF-based communication can provide better connectivity.

Hybrid Systems in Underwater Communication

Underwater wireless communication mainly relies on acoustic, RF, and optical means [54]. The advantages and disadvantages of these three underwater wireless communication technologies are compared in Table 46.1. From the table, it can be seen that underwater optical communication has high speed, high confidentiality,

Table 46.1 Differences of three underwater wireless communication technologies

Parameter	RF	Acoustic	Underwater optical
Frequency	30 Hz – 300 Hz or MHz	10 Hz – 100 kHz	10^{12} Hz – 10^{15} Hz
Transmit power	Milliwatt to watt range	Tens of watts	Few watts
Data rate	Mbit/s	Kbit/s	Gbit/s
Transmission distance	$\leq$10 m	$\leq$20 km	10 – 100 m
Performance parameters	Conductivity and dielectric constant	Temperature, salinity	Absorption, scattering, turbidity, and organic matter
Transmission delay	Medium	High	Low
Attenuation	3.5 – 5 dB/m	0.1 – 4 dB/km	0.39 dB/m (attenuation in turbid waters is 11 dB/m)
LOS/NLOS	Both	Both	Only LOS

low latency, and low cost but cannot be used for nonline-of-sight transmissions. The system performance is greatly affected by seawater absorption and scattering, channel turbulence, etc., which is complementary to RF and underwater acoustic communication. Therefore, using two or more underwater communication systems to form a hybrid system can overcome the limitations of a single communication system.

Mohsan et al. [55] investigated a hybrid underwater acoustic/optical wireless communication system for real-time video monitoring between an autonomous underwater vehicle and a base station. The system was equipped with optical and acoustic modems to exploit the advantages of both communication methods. An additional acoustic receiver was equipped to improve communication redundancy, while providing additional data to locate another transmitter and align the optical modem.

During the alignment process, data is sent over the acoustic link until the optical modem is precisely aligned. When the optical modem is aligned, video data is sent using the optical link, thus increasing the system bandwidth. Even when the optical link is not available, the acoustic link remains operational to exchange control signals and send video frames. This hybrid acoustic/optical system can be dynamically configured to maintain a high-speed optical link.

Naik et al. [56] introduced the concept of a reconfigurable intelligent surface to underwater wireless communication systems due to their ability to improve the energy and spectral efficiency of wireless communication links by reconfiguring the propagation environment of radio frequency/optical waves. The reconfigurable intelligent surface is a full duplex relay without amplification. It has a large number of passive reflective elements integrated into its plane to reflect an incident signal in the desired direction and outperforms conventional relay systems (such as amplify-and-forward and decode-and-forward) in terms of communication rate and

complexity. Placing M reconfigurable intelligent surfaces at different locations with sunken objects underwater solves the optical link problem of nonline-of-sight transmissions due to underwater obstacle occlusion. Increasing the number of reconfigurable intelligent surfaces can improve the average signal-to-noise ratio of the underwater OWC/RF hybrid system.

Hybrid Systems in 5G Retransmission

In 5G communication, ultradense and high-traffic cells are connected to a core network through backhaul links; however, they typically have extremely high requirements for capacity, latency, availability, and cost. To solve the 5G backhaul problem, Jaber et al. [57] proposed a 5G cloud radio access network (C-RAN), whose network architecture is shown in Fig. 46.9.

In a 5G C-RAN uplink, users transmit data to a remote radio unit (RRU) through an RF link. The RRU acts as a relay and forwards the received data to a baseband unit (BBU) through an OWC link. This process is defined as a 5G front-end link. The C-RAN architecture utilizes the diversity of traffic peaks to improve the infrastructure utilization and promote the green development of 5G.

The Raytheon Corporation of the United States and NxGen Corporation [58] have developed a low-cost, high-bandwidth 5G backhaul application. It uses multiple orbital-angular-momentum (OAM) modes to multiplex and demultiplex orthogonal and independent optical channels to improve the link throughput and isolate the channels from RF full duplex backhaul links. Using the OAM method to coaxially encapsulate RF and OWC links provides a high-frequency spectral efficiency wireless backhaul link. Compared to optical fiber backhaul, wireless backhaul links based on hybrid OWC/RF configurations have lower latency and can function

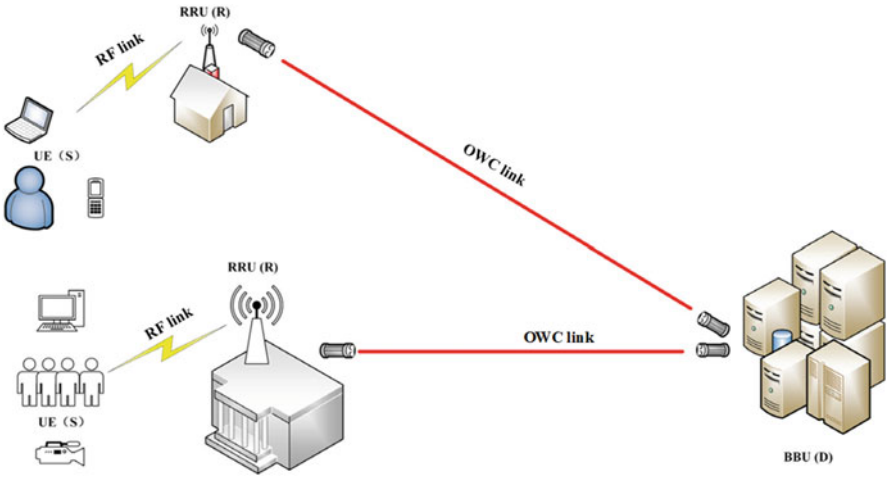


Fig. 46.9 OWC/RF hybrid communication in a 5G cloud wireless access network [57]

even in severe weather, thus meeting the availability and reliability requirements of modern telecommunication networks.

Hybrid System Performance

OWC Subsystem

In an OWC/RF hybrid parallel communication system, the instantaneous signal-to-noise ratio (SNR) of the OWC link can be expressed as [59]

$$\gamma^{\mathrm{OWC}} = \frac{\left(P^{\mathrm{OWC}}\eta^{\mathrm{OWC}}R\xi\right)^2}{\sigma_n^{\mathrm{OWC},2}}\left(h^{\mathrm{OWC}}\right)^2 = \overline{\gamma}^{\mathrm{OWC}}\left(h^{\mathrm{OWC}}\right)^2, \tag{46.1}$$

where P^{OWC} denotes the transmit power of the OWC link, η^{OWC} denotes the atmospheric loss, R denotes the response of the photodetector, ξ denotes the modulation index, $\sigma_n^{\mathrm{OWC},2}$ denotes the noise variance in the OWC link, h^{OWC} denotes the channel gain of the OWC link, and $\overline{\gamma}^{\mathrm{OWC}}$ denotes the average signal-to-noise ratio of the OWC link.

Assuming that the OWC channel obeys an exponential Weibull distribution, the probability density function corresponding to the instantaneous signal-to-noise ratio γ^{FSO} of the OWC link [60] is

$$f_{\gamma^{\mathrm{OWC}}}\left(\gamma^{\mathrm{OWC}}\right) = \frac{1}{2\sqrt{\gamma^{\mathrm{OWC}}\overline{\gamma}^{\mathrm{OWC}}}}\frac{\alpha\rho^2}{(\eta A_0)^{\rho^2}}\left(\sqrt{\frac{\gamma^{\mathrm{OWC}}}{\overline{\gamma}^{\mathrm{OWC}}}}\right)^{\rho^2-1} \cdot \sum_{t=0}^{\infty} B_t G_{1,2}^{2,0}\left[C_t\left(\sqrt{\frac{\gamma^{\mathrm{OWC}}}{\overline{\gamma}^{\mathrm{OWC}}}}\right)^{\beta}\middle|\begin{matrix}1\\0,1-\frac{\rho^2}{\beta}\end{matrix}\right], \tag{46.2}$$

where ρ is the ratio of the received beam radius to the standard deviation of the jitter at the receiving plane, α and β are the shape parameters of the turbulence, η is the scale parameter, $G_{p,q}^{m,n}(\cdot)$ denotes the Meijer G-function, $B_t = \frac{(-1)^t\Gamma(\alpha)}{t!\Gamma(\alpha-t)(1+t)^{1-\frac{\rho^2}{\beta}}}$, $C_t = \frac{(1+t)}{(\eta A_0)^{\beta}}$, $\Gamma(\cdot)$ denotes the gamma function, $A_0 = \mathrm{erf}^2(v)$, where $\mathrm{erf}(\cdot)$ is the complementary error function, $v = \sqrt{\frac{\pi}{2}}\frac{R_a}{w_b}$, R_a is the radius of the receiver aperture, and w_b is the beam radius at transmission distance L.

The cumulative distribution function of γ^{OWC} obtained by integrating Eq. (46.2) is

$$F_{\gamma^{\mathrm{OWC}}}\left(\gamma^{\mathrm{OWC}}\right) = A\left(\sqrt{\frac{\gamma^{\mathrm{OWC}}}{\overline{\gamma}^{\mathrm{OWC}}}}\right)^{\rho^2}\sum_{t=0}^{\infty} B_t G_{2,3}^{2,1}\left[C_t\left(\sqrt{\frac{\gamma^{\mathrm{OWC}}}{\overline{\gamma}^{\mathrm{OWC}}}}\right)^{\beta}\middle|\begin{matrix}\kappa_1\\\kappa_2\end{matrix}\right]. \tag{46.3}$$

RF Subsystem

The instantaneous signal-to-noise ratio of the RF link in the OWC/RF hybrid parallel communication system [29] is expressed as

$$\gamma^{\mathrm{RF}} = \frac{P^{\mathrm{RF}}\eta^{\mathrm{RF}}\left(h^{\mathrm{RF}}\right)^2}{\sigma_n^{\mathrm{RF},2}} = \overline{\gamma}^{\mathrm{RF}}\left(h^{\mathrm{RF}}\right)^2, \tag{46.4}$$

where P^{RF} denotes the transmit power of the RF link, η^{RF} denotes the RF channel loss, h^{RF} denotes the channel state, $\sigma_n^{\mathrm{RF},2}$ denotes the noise variance of the RF link, and $\overline{\gamma}^{\mathrm{RF}}$ denotes the average signal-to-noise ratio of the RF link.

Assuming that the RF channel follows a Nakagami-m distribution, the probability density function corresponding to the instantaneous signal-to-noise ratio γ^{RF} of the RF link [61] is

$$f_{\gamma\mathrm{RF}}\left(\gamma^{\mathrm{RF}}\right) = \left(\frac{m}{\overline{\gamma}^{\mathrm{RF}}}\right)^m \frac{\gamma^{m-1}}{\Gamma(m)} G_{0,1}^{1,0}\left[\frac{m\gamma^{\mathrm{RF}}}{\overline{\gamma}^{\mathrm{RF}}}\middle|\begin{matrix}-\\0\end{matrix}\right], \tag{46.5}$$

where m denotes the parameters of the Nakagami-m distribution, $\Gamma(\cdot)$ denotes the gamma function, and $G_{p,q}^{m,n}(\cdot)$ denotes the Meijer G-function.

The cumulative distribution function of γ^{RF} obtained by integrating Eq. (46.5) is

$$F_{\gamma\mathrm{RF}}\left(\gamma^{\mathrm{RF}}\right) = \frac{1}{\Gamma(m)} G_{1,2}^{1,1}\left[\frac{m\gamma^{\mathrm{RF}}}{\overline{\gamma}^{\mathrm{RF}}}\middle|\begin{matrix}1\\m,0\end{matrix}\right]. \tag{46.6}$$

Hybrid System

Signal-to-Noise Ratio

If a selective combining scheme is adopted at the receiving end, the output signal-to-noise ratio [62] of the hybrid OWC/RF parallel communication system is

$$\gamma = \max\left(\gamma^{\mathrm{OWC}}, \gamma^{\mathrm{RF}}\right). \tag{46.7}$$

The cumulative distribution function [62] corresponding to the instantaneous signal-to-noise ratio γ of the hybrid system is

$$\begin{aligned} F_\gamma(\gamma) &= F_{\gamma\mathrm{RF}}\left(\gamma^{\mathrm{RF}}\right) F_{\gamma\mathrm{OWC}}\left(\gamma^{\mathrm{OWC}}\right) \\ &= \frac{A}{\Gamma(m)}\left(\frac{\gamma}{\overline{\gamma}^{\mathrm{OWC}}}\right)^{\frac{\rho^2}{2}} \sum_{t=0}^{\infty} B_t G_{0,0;\,1,1;\,2,1}^{0,0;\,1,1;\,2,1}\left[\begin{matrix}-\\-\end{matrix}\middle|\begin{matrix}1\\m,0\end{matrix}\middle|\begin{matrix}\kappa_1\\\kappa_2\end{matrix}\middle|\frac{m\gamma}{\overline{\gamma}^{\mathrm{RF}}}, C_t\left(\frac{\gamma}{\overline{\gamma}^{\mathrm{OWC}}}\right)^{\frac{\beta}{2}}\right]. \end{aligned} \tag{46.8}$$

Here, G is the extended generalized bivariate Meiger G-function and γ is the output signal-to-noise ratio of the hybrid parallel system.

Outage Probability

When the instantaneous signal-to-noise ratio falls below the set threshold, the system is disrupted and cannot communicate. The disruption probability of the hybrid system is expressed as [29]

$$P_{\text{out}} = \Pr\{\gamma < \gamma_{\text{th}}\} = F_{\gamma}(\gamma_{\text{th}}) = \frac{A}{\Gamma(m)}\left(\frac{\gamma_{\text{th}}}{\overline{\gamma}^{\text{OWC}}}\right)^{\frac{\rho^2}{2}} \sum_{t=0}^{\infty} B_t G_{0,0;\,1,2;\,2,3}^{0,0;\,1,1;\,2,1}\left[\begin{matrix} - \\ - \end{matrix}\middle|\begin{matrix} 1 \\ m,0 \end{matrix}\middle|\begin{matrix} \kappa_1 \\ \kappa_2 \end{matrix}\middle|\frac{m\gamma_{\text{th}}}{\overline{\gamma}^{\text{RF}}}, C_t\left(\frac{\gamma_{\text{th}}}{\overline{\gamma}^{\text{OWC}}}\right)^{\frac{\beta}{2}}\right], \tag{46.9}$$

where $\Pr\{\cdot\}$ represents the probability, $F_{\gamma}(\cdot)$ represents the cumulative distribution function of the signal-to-noise ratio, $\Gamma(\cdot)$ represents the gamma function, γ_{th} represents the set threshold, ρ is the ratio of the received-beam radius to the jitter standard deviation of the receiving plane, and $\overline{\gamma}^{\text{OWC}}$ $\overline{\gamma}^{\text{RF}}$ represent the average signal-to-noise ratio of the OWC and RF links, respectively. G is the extended generalized bivariate Meiger G-function, β is the shape parameter of the turbulence, $B_t = \frac{(-1)^t\Gamma(\alpha)}{t!\Gamma(\alpha-t)(1+t)^{1-\frac{\rho^2}{\beta}}}$, $C_t = \frac{(1+t)}{(\eta A_0)^{\beta}}$, and m denotes the distribution parameters of the RF channel.

Average Bit Error Rate

The average bit error rate refers to the ratio of the number of error symbols to the total number of symbols in the transmission. It is an important indicator to evaluate the performance of a communication system. If the binary modulation scheme is adopted, the average bit error rate of the hybrid system [63] is expressed as follows:

$$P_b = \frac{Aq^p}{2\Gamma(p)\Gamma(m)}\left(\overline{\gamma}^{\text{OWC}}\right)^{-\frac{\rho^2}{2}} \sum_{t=0}^{\infty} B_t \sum_{\tau=1}^{n} \omega_{\tau} \cdot G_{0,0;\,1,2;\,2,3}^{0,0;\,1,1;\,2,1}\left[\begin{matrix} - \\ - \end{matrix}\middle|\begin{matrix} 1 \\ m,0 \end{matrix}\middle|\begin{matrix} \kappa_1 \\ \kappa_2 \end{matrix}\middle|\frac{mx_{\tau}}{\overline{\gamma}^{\text{RF}}}, C_t\left(\frac{x_{\tau}}{\overline{\gamma}^{\text{OWC}}}\right)^{\frac{\beta}{2}}\right], \tag{46.10}$$

where p and q are the bit error rate parameters describing different binary modulation schemes, ω_{τ} is the weight functions, and x_{τ} is the coordinate special point.

Problems, Challenges, and Solutions

Weather Conditions

Table 46.2 shows the impact of different weather conditions on communication links, taking 40-GHz RF/OWC hybrid communication as an example [5]. As can be seen from Table 46.2, OWC links are greatly affected by weather. In snow, fog, and other weather conditions, the availability and reliability of OWC links cannot be guaranteed, while RF links are very sensitive to rain. In this case, hybrid communication becomes a feasible solution to these challenges.

Table 46.2 [5] Availability of 40-GHz RF and OWC links in different weather conditions

Weather conditions	RF (40 GHz)	OWC	Hybrid
Snow	100%	39.49%	100%
Dense fog	100%	0.51%	100%
Rain	14.29%	85.71%	85.71

Frequency selection is crucial for hybrid systems [63]. An incorrect frequency can lead to severe signal attenuation and a significant reduction in transmission distance. If the frequency is too low and the signal wavelength is long, diffraction phenomena and interference are prone to occur. The frequency of RF links is closely related to weather conditions, communication rates, and other factors. By understanding link attenuation and selecting an appropriate frequency, the impact of weather on the performance of hybrid OWC/RF communication systems can be effectively reduced.

Atmospheric Turbulence

In hybrid communication systems, laser beams generated by OWC links and transmitted through atmospheric channels are affected by atmospheric turbulence. This can lead to phenomena, such as light-intensity flicker, beam drift, and expansion [65, 66], which reduce the reliability of the hybrid system. To address the adverse effects caused by turbulence and ensure the normal operation of the OWC link in a hybrid communication system, the following measures can be taken:

1. For short-range communication, adaptive optic (AO) technology is generally used to compensate [67]. Owing to factors such as the antenna aperture and link bandwidth, the operating range of an AO system is within the Fresnel region. When this region is exceeded, the resulting aberration cannot be adaptively compensated by the antennas at the receiving and transmitting ends.
2. For long-distance communication, automatic optical gain control technology can be used to compensate [68]. At the receiving end, the gain can be adjusted through a feedback circuit to maintain the stability of the output signal and improve the system performance.

Pointing Errors

The OWC link requires a strict alignment between the optical transmitter and optical receiver to establish a reliable connection. However, in practice, owing to factors such as wind load, thermal expansion, and building shaking, laser signals deviate from their original direction, and beam-pointing errors occur between transceivers [69]. Because of the limited receiving aperture, if the pointing errors are too large, the receiving end cannot detect the optical signal, resulting in a communication interruption.

To mitigate the adverse impact of pointing errors on OWC performance [46], beamwidth optimization can be used to control the beamwidth through the

transmission aperture of the OWC system to minimize the geometric loss and error probability. In addition, because RF links are not prone to pointing errors, millimeter-wave RF links can be used as the main data transmission link to reduce the impact of pointing errors on hybrid communication systems; however, in this case, the system communication data rate will also be reduced.

Multipath Effect

In RF communication links, when electromagnetic waves propagate in the channel, they are reflected by objects, causing signals to arrive at the receiver from different paths, producing a multipath effect [70]. Owing to the different propagation paths, the time taken to reach the receiver varies. Each component overlaps with the others, according to its respective phase, causing interference, which leads to signal distortion. Multipath effects can also lead to signal fading.

A decision feedback equalizer can reduce the impact of multipath effects on hybrid communication systems [71]. The equalizer is mainly composed of feedforward and feedback filters. It is a limited-length rake receiver that uses the ratio of signal-to-interference plus noise to measure the signal quality. By eliminating linear interference in the feedforward portion, it can eliminate the intersymbol interference caused by subsequent multipath delays.

Multipath diversity reception technology can distinguish subtle multipath signals in real time, adjust the weights of the identified multipath signals, and composite them into enhanced signals. This not only solves interference problems but also enhances effective signals.

Summary

As an important research direction in the field of optical wireless communication, hybrid OWC/RF communication combines the advantages of both OWC and RF communication methods. This not only overcomes the OWC links' susceptibility to weather but also addresses the limited RF link communication rate; it improves the system reliability while considering the communication rate. Providing solutions for the "last mile" access of mobile communications has broad application prospects in underwater, unmanned aerial vehicles, satellites, and 5G communications.

This chapter described the development of hybrid OWC/RF communication systems over the past 15 years, introduced two main hybrid communication system models, discussed hybrid communication systems in different application scenarios, analyzed the problems and challenges faced by hybrid communication, presented the key technologies to solve these problems, and projected the development trends of hybrid OWC/RF communication [46, 48, 52]. These mainly include the following aspects:

1. Improving the adaptive-switching algorithm of hybrid OWC/RF parallel communication systems to optimize the hybrid system performance, while reducing the number of link switches.
2. In view of the wireless channel asymmetry in actual hybrid communication, studying hybrid modulation and coding techniques [46, 72, 73] can improve the reliability and resource utilization (power, bandwidth, budget, etc.) of hybrid systems.
3. Researching hybrid communication systems in spatial information networks [74] to improve the coverage and real-time information processing capabilities of spatial information networks. Building an interconnected space–space integrated information communication network and maximizing the utilization of spatial information resources.
4. Studying the hybrid communication system in a 5G backhaul design [75] to solve network congestion, latency, and other issues and thereby accelerating the construction of 5G+ and 6G communication systems.

References

1. Hamza A S, Deogun J S, Alexander D R. Classification framework for free space optical communication links and systems [J]. IEEE Communications Surveys & Tutorials, 2018, 21(2): 1346–1382.
2. Khalighi M A, Uysal M. Survey on free space optical communication: A communication theory perspective [J]. IEEE communications surveys & tutorials, 2014, 16(4): 2231–2258.
3. Khalid H, Muhammad S, Nistazakis H E, et al. Performance analysis of hard-switching based hybrid FSO/RF system over turbulence channels [J]. Computation, 2019, 7(2): 28–38.
4. He B, Schober R. Bit-interleaved coded modulation for hybrid RF/FSO systems [J]. IEEE Transactions on Communications, 2009, 57(12): 3753–3763.
5. Nadeem F, Kvicera V, Awan M S, et al. Weather effects on hybrid FSO/RF communication link [J]. IEEE Journal on Selected Areas in Communications, 2009, 27(9): 1687–1697.
6. Abdulhussein A, Oka A, Nguyen T T, et al. Rateless coding for hybrid free-space optical and radio-frequency communication [J]. IEEE Transactions on Wireless Communications, 2010, 9(3): 907–913.
7. Lee E, Park J, Han D, et al. Performance analysis of the asymmetric dual-hop relay transmission with mixed RF/FSO links [J]. IEEE Photonics Technology Letters, 2011, 23(21): 1642–1644.
8. Ansari I S, Yilmaz F, Alouini M S. Impact of pointing errors on the performance of mixed RF/FSO dual-hop transmission systems [J]. IEEE Wireless Communications Letters, 2013, 2(3): 351–354.
9. Usman M, Yang H C, Alouini M S. Practical switching-based hybrid FSO/RF transmission and its performance analysis [J]. IEEE Photonics Journal, 2014, 6(5): 1–13.
10. Touati A, Abdaoui A, Touati F, et al. On the effects of combined atmospheric fading and misalignment on the hybrid FSO/RF transmission [J]. Journal of Optical Communications and Networking, 2016, 8(10): 715–725.
11. Yang L, Hasna M O, Ansari I S. Unified performance analysis for multiuser mixed η-μ and M-distribution dual-hop RF/FSO systems [J]. IEEE Transactions on Communications, 2017, 65(8): 3601–3613.
12. Amirabadi M A, Vakili V T. A novel hybrid FSO/RF communication system with receive diversity [J]. Optik, 2019, 184(3): 293–298.

13. Esmail M A, Ragheb A M, Fathallah H A, et al. 5G-28 GHz signal transmission over hybrid all-optical FSO/RF link in dusty weather conditions [J]. IEEE Access, 2019, 7(2): 24404–24410.
14. Ashrafzadeh B, Soleimani N E, Zaimbashi A, et al. Outage performance of mixed RF-FSO systems over DGG and Nakagami-m channels [J]. IEEE Wireless Communications Letters, 2020, 9(12): 2135–2139.
15. Khan M N, Kashif H, Rafay A. Performance and optimization of hybrid FSO/RF communication system in varying weather [J]. Photonic Network Communications, 2021, 41(1): 47–56.
16. Vishwakarma N, Swaminathan R. On the capacity performance of hybrid FSO/RF system with adaptive combining over generalized distributions [J]. IEEE Photonics Journal, 2022, 14(1): 1–12.
17. Magidi S, Jabeena A. Analysis of hybrid FSO/RF communication system under the effects of combined atmospheric fading and pointing errors [J]. Optical and Quantum Electronics, 2022, 54(4): 1–20.
18. Shu Fang, Ao Faliang, Liao Xinding. Research on automatic switching conditions of hybrid FSO/RF system [J]. Journal Of Guilin University Of Electronic Technology, 2008, 94(01): 1–4.
19. Meng Wen. Wireless laser/microwave dual-link communication control technology research [D]. Xi'an: Xidian University, 2009.
20. Su Jia, Chen Hui. Research on hybrid FSO/RF system switching technology based on network packet loss rate [J]. Optical Communication Technique, 2011, 35(10): 28–38.
21. Yang L, Hasna M O, Gao X. Performance of mixed RF/FSO with variable gain over generalized atmospheric turbulence channels [J]. IEEE Journal on Selected Areas in Communications, 2015, 33(9): 1913–1924.
22. Chen L, Wang W, Zhang C. Multiuser diversity over parallel and hybrid FSO/RF links and its performance analysis [J]. IEEE Photonics Journal, 2016, 8(3): 1–9.
23. Zhao Jing, Zhao Shanghong, Zhao Weihu, et al. Performance analysis of hybrid RF/FSO aerial communication systems under atmospheric turbulence and pointing errors [J]. Chinese Journal of Lasers, 2017, 44(09): 176–184.
24. Zhao J, Zhao S H, Zhao W H, et al. Performance of mixed RF/FSO systems in exponentiated Weibull distributed channels [J]. Optics Communications, 2017, 405(7): 244–252.
25. Shao Junhu, Su Runmin, Yao Liu, et al. Performance analysis of soft-switching FSO/RF link hybrid coding and modulation algorithm [J]. Journal of Electronic Measurement and Instrument, 2017, 31(05): 682–687.
26. Yi X, Shen C, Yue P, et al. Performance of decode-and-forward mixed RF/FSO system over $\kappa-\mu$ shadowed and exponentiated Weibull fading [J]. Optics Communications, 2019, 439(1): 103–111.
27. Liu Hongzhan, Jiang Ting, Hao Yuan. Research on improving deep space communication using hybrid RF-FSO system [J]. Journal of Communication, 2020, 41(10): 148–155.
28. Liu Wenya, Zhao Shanghong, Wang Xiang, et al. Performance analysis of hybrid FSO/RF aerial communication system based on SC diversity [J]. Laser and Infrared, 2020, 50(05): 525–538.
29. Wu Yan, Mei Haiping, Dai Congming, et al. Performance analysis of hybrid free space optical/RF communication system [J]. Advances in Lasers and Optoelectronics, 2021, 58(17): 203–209.
30. Liu Wei, Wang Jingyi, Ma Xinyang, et al. Hybrid millimeter-wave radio frequency/multiple input multiple output free space optical communication relay system [J]. Optics and Precision Engineering, 2022, 30(05): 536–544.
31. Ding Liang, Wu Zhiyong, Gu Yucong, et al. Key technologies for wireless laser and RF complementary communication systems [J]. Advances in Lasers and Optoelectronics, 2019, 56(06): 49–64.
32. Wu Y, Yang Q, Park D, et al. Dynamic link selection and power allocation with reliability guarantees for hybrid FSO/RF systems [J]. IEEE Access, 2017, 5: 13654–13664.
33. Mai V, Pham A T. Adaptive multi-rate designs for hybrid FSO/RF systems over fading channels [C] //2014 IEEE Globecom Workshops. Austin: IEEE, 2014: 469–474.

34. Nadeem F, Geiger B, Leitgeb E, et al. Comparison of link selection algorithms for free space optics/radio frequency hybrid network [J]. IET Communications, 2011, 5(18): 2751–2759.
35. Vishwakarma N, Swaminathan R. Performance analysis of hybrid FSO/RF communication over generalized fading models [J]. Optics Communications, 2021, 487: 126796–126813.
36. Zhang W, Hranilovic S, Shi C. Soft-switching hybrid FSO/RF links using short-length raptor codes: design and implementation [J]. IEEE Journal on Selected Areas in Communications, 2009, 27(9): 1698–1708.
37. Eslami A, Vangala S, Pishro-Nik H. Hybrid channel codes for efficient FSO/RF communication systems [J]. IEEE Transactions on Communications, 2010, 58(10): 2926–2938.
38. Bag B, Das A, Ansari I S, et al. Performance analysis of hybrid FSO systems using FSO/RF-FSO link adaptation [J]. IEEE Photonics Journal, 2018, 10(3): 1–17.
39. Miridakis N I, Matthaiou M, Karagiannidis G K. Multiuser relaying over mixed RF/FSO links [J]. IEEE Transactions on Communications, 2014, 62(5): 1634–1645.
40. Wang Y, Wang P, Liu X, et al. On the performance of dual-hop mixed RF/FSO wireless communication system in urban area over aggregated exponentiated Weibull fading channels with pointing errors [J]. Optics Communications, 2018, 410: 609–616.
41. Han Liqiang, Jiang Hongbing. Performance analysis of a two-hop hybrid RF/MIMO FSO amplify-and-forward relay communication system [J]. Advances in Lasers and Optoelectronics, 2020, 57(13): 1–10.
42. AO Qingqing. Performance study of multi-user dual-hop hybrid RF/FSO system [D]. Xi'an: Xidian University, 2021.
43. Jamali V, Michalopoulos D S, Uysal M, et al. Link allocation for multiuser systems with hybrid RF/FSO backhaul: Delay-limited and delay-tolerant designs[J]. IEEE Transactions on Wireless Communications, 2016, 15(5): 3281–3295.
44. Najafi M, Schmauss B, Schober R. Intelligent reflecting surfaces for free space optical communication systems [J]. IEEE Transactions on Communications, 2021, 69(9): 6134–6151.
45. Yang L, Guo W, Ansari I S. Mixed dual-hop FSO-RF communication systems through reconfigurable intelligent surface [J]. IEEE Communications Letters, 2020, 24(7): 1558–1562.
46. Stotts L B, Andrews L C, Cherry P C, et al. Hybrid optical RF airborne communications [J]. Proceedings of the IEEE, 2009, 97(6): 1109–1127.
47. Rödiger B, Ginthör D, Labrador J P, et al. Demonstration of an FSO/RF hybrid-communication system on aeronautical and space applications [C] //Conference on Laser Communication and Propagation through the Atmosphere and Oceans. United States: SPIE, 2020: 1–11.
48. Mohsan S A H, Amjad H. A comprehensive survey on hybrid wireless networks: practical considerations, challenges, applications and research directions [J]. Optical and Quantum Electronics, 2021, 53(9): 1–56.
49. Liang Jingyuan, Kang Weilong, Dong Zhuang, et al. Research progress of optical antenna technology for free-space optical communication system [J]. Optical Communication Technology, 2022, 46(04): 1–10.
50. Ke Xizheng, Qin Huanhuan, Yang Shangjun, et al. Nighttime background light noise model for vehicle network visible light communication system [J]. Chinese Journal. of Radio Science, 2021, 36(06): 986–990.
51. Nguyen T K, Nguyen C T, Le H D, et al. TCP performance over satellite-based hybrid FSO/RF vehicular networks: Modeling and analysis [J]. IEEE Access, 2021, 9: 108426–108440.
52. Zhou K, Gong C, Wu N, et al. Distributed channel allocation and rate control for hybrid FSO/RF vehicular ad hoc networks [J]. Journal of Optical Communications and Networking, 2017, 9(8): 669–681.
53. Chowdhury M Z, Hossan M T, Hasan M K, et al. Integrated RF/optical wireless networks for improving QoS in indoor and transportation applications [J]. Wireless Personal Communications, 2019, 107(3): 1401–1437.
54. Chowdhury M Z, Hasan M K, Shahjalal M, et al. Optical wireless hybrid networks: Trends, opportunities, challenges, and research directions [J]. IEEE Communications Surveys & Tutorials, 2020, 22(2): 930–966.

55. Mohsan S A H, Hasan M, Mazinani A, et al. A systematic review on practical considerations, recent advances and research challenges in underwater optical wireless communication [J]. International Journal of Advanced Computer Science and Applications, 2020, 11(7): 1–11.
56. Naik R P, Chung W Y. Evaluation of reconfigurable intelligent surface-assisted underwater Optical Wireless Communication system [J]. Journal of Lightwave Technology, 2022, 40(13): 4257–4267.
57. Jaber M, Imran M A, Tafazolli R, et al. 5G backhaul challenges and emerging research directions: A survey [J]. IEEE Access, 2016, 4: 1743–1766.
58. Graceffo G M, Talamonti J, Campbell L, et al. Hybrid RF & FSO for defense and 5g backhaul [C]//2019 IEEE Globecom Workshops. Waikoloa: IEEE, 2019: 1–6.
59. Song X G, Yang F, Cheng J L. Subcarrier intensity modulated optical wireless communications in atmospheric turbulence with pointing errors [J]. Journal of Optical Communications and Networking, 2013, 5(4): 349–358.
60. Jing Z, Zhao S H, Zhao W H, et al. Performance analysis for mixed FSO/RF Nakagami-m and Exponentiated Weibull dual-hop airborne systems [J]. Optics Communications, 2017, 392: 294–299.
61. Alathwary W A, Altubaishi E S. On the performance analysis of decode-and-forward multi-hop hybrid FSO/RF systems with hard-switching configuration [J]. IEEE Photonics Journal, 2019, 11(6): 1–12.
62. Shakir W M R. Performance evaluation of a selection combining scheme for the hybrid FSO/RF system [J]. IEEE Photonics Journal, 2017, 10(1): 1–10.
63. Ansari I S, Al-Ahmadi S, Yilmaz F, et al. A new formula for the BER of binary modulations with dual-branch selection over generalized-K composite fading channels [J]. IEEE Transactions on Communications, 2011, 59(10): 2654–2658.
64. Oguchi T. Electromagnetic wave propagation and scattering in rain and other hydrometeors [J]. Proceedings of the IEEE, 1983, 71(9): 1029–1078.
65. Li Zheng, Liao Zhiwen, Liang Jingyuan, et al. Research and prospects of atmospheric turbulence model [J]. Optical Communication Technology, 2022, 46(12): 1–13.
66. Ke Xizheng, Wu Jiali, Yang Shangjun. Progress and prospects of atmospheric turbulence research for Optical Wireless Communication [J]. Chinese Journal. of Radio Science, 2021, 36(03): 323–339.
67. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress of adaptive optics for Optical Wireless Communication system at Xi'an University of Technology [J]. Intense Lasers and Particle Beams, 2021, 33(08): 30–52.
68. Sharma S, Madhukumar A S, Swaminathan R. Effect of pointing errors on the performance of hybrid FSO/RF networks [J]. IEEE Access, 2019, 7: 131418–131446.
69. Mohsan S A H, Khan M A, Amjad H. Hybrid FSO/RF networks: A review of practical constraints, applications and challenges [J]. Optical Switching and Networking, 2022, 47: 100697–100713.
70. Ke Xizheng, Deng Lijun. Optical wireless communication [M]. Beijing: Science Press, 2016.
71. Gao Xiaomei, Shu Yuting, Liang Jingyuan, et al. Research progress of communication laser and its modulation technology [J]. Study on Optical Communications, 2022, 48(11): 1–19.
72. Ke Xizheng, Liang Jingyuan, Xu Dongsheng, et al. Research progress of pulse position modulation technology for Optical Wireless Communication class [J]. Photovoltaic Engineering, 2022, 49(03): 3–21.
73. Vangala S, Pishro-Nik H. Hybrid channel codes for highly efficient FSO/RF communication systems [J]. IEEE Transactions on Wireless Communications, 2010, 58(10): 2926–2938.
74. Xu G, Song Z. Effects of solar scintillation on deep space communications: challenges and prediction techniques [J]. IEEE Wireless Communications, 2019, 26(2): 10–16.
75. Baha M A M, Alsaify A, Alma'aitah A Y. Innovative unmanned aerial vehicle self-backhauling hybrid solution using RF/FSO system for 5G networks [J]. International Journal of Electrical & Computer Engineering, 2022, 12(4): 4483–4506.

Performance Analysis of a Hybrid FSO/RF/THz Relay Communication System

47

Contents

Introduction 1730
Foreign and Domestic Research Progress 1731
System and Channel Models 1735
Channel Model 1735
FSO Link 1736
RF Link 1737
THz Link 1738
Performance Analysis 1739
Outage Probability 1739
Average Bit Error Rate 1741
Simulation Results 1743
Conclusion 1750
References 1751

Abstract

A dual-hop hybrid free-space optical/radio frequency (FSO/RF) relay system is considered an effective solution for increasing the capacity and coverage of wireless communication systems between last-kilometer access networks and backbone networks. However, the transmission performance of FSO links is greatly affected by atmospheric turbulence, pointing errors, fog, and snow. The RF access network has a low rate and high delay. To ensure communication quality, a hybrid FSO/RF/terahertz (THz) relay system was designed. Decode-and-forward relaying was selected and hybrid FSO/RF communication based on an adaptive-combining scheme was considered before the relay nodes. THz links were used after the relay nodes to reach the users. Using the statistical characteristics of the different obtained links, expressions for the outage probability and average bit error rate of the system were derived. The effects of different parameters on the performance of hybrid FSO/RF/THz relay systems were studied, and the first-hop adaptive-combining scheme was compared with the single-threshold switching scheme and single-link FSO system. The analysis

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_47

results indicated that the hybrid FSO/RF/THz relay system provided better outage and error performance using an adaptive-combining scheme. It improved the system communication rate while considering the communication reliability.

Keywords

Free-space optical (FSO) communication · Radio frequency (RF) communication · Terahertz (THz) communication · Adaptive-combining scheme

Introduction

Free-space optical (FSO) communications have received considerable attention in the field of wireless communications, owing to their advantages of a high-speed rate, large capacity, high confidentiality, low cost, and ease of deployment. They are the primary choice for connecting radio frequency (RF) access networks and backbone networks.

However, FSO is highly sensitive to problems, such as atmospheric turbulence, pointing errors, and weather conditions (e.g., fog, haze, and snow), which limit the transmission distance and seriously affect the performance of FSO communications [1–3]. On the other hand, RF communication only attenuates severely in rain, is insensitive to atmospheric turbulence and pointing errors, and has strong environmental adaptability. However, RF communication has disadvantages, such as poor interference resistance, low security, and a low transmission rate [4].

In addition, terahertz (THz) communication (0.1–10 THz) is one of the core technologies of B5G and 6G communication and has a frequency band between RF and FSO. It has the advantages of low latency, high directional gain, and high available bandwidth. However, the huge path loss, molecular absorption, pointing errors, and other problems lead to a limited transmission distance; thus, THz is only suitable for short-distance high-rate communication [5, 6].

Hybrid communication combines the respective advantages and disadvantages of FSO, RF, and THz communications. Hybrid communication refers to the use of two or more different communication modes to transmit data or information in a communication system. It can overcome the limitations of a single communication mode and access the complementary advantages of the others. Currently, research on hybrid communication systems mainly focuses on hybrid FSO/RF, FSO/THz, and THz/RF communication systems; few reports exist on communication systems with a mixture of FSO, RF, and THz.

In this study, a hybrid FSO/RF relay system was improved to form a hybrid FSO/RF/THz relay communication system. Before the relay node, a hybrid FSO/RF communication system, based on an adaptive-combining scheme, was used to alleviate the degradation of the system performance of the separate FSO link caused by atmospheric turbulence and other factors. After the relay node, the THz link was utilized instead of the traditional millimeter-wave RF link to access the user, thus

further improving the system communication rate. The outage probability and average bit error rate expressions of the system were derived. The correctness of the derived performance formula was verified by simulating the relationship between different parameters and the performance of the hybrid FSO/RF/THz relay system.

Foreign and Domestic Research Progress

In recent years, many experts and scholars have extensively researched different types of hybrid systems, considering their advantages in outdoor wireless communications, such as high speed, high reliability, long distance, and extensive coverage.

In the early stages of hybrid communication system research, owing to the complementary characteristics of FSO and RF links, research mainly focused on hybrid FSO/RF communication systems. However, in recent years, as research on THz communication has deepened, THz technology has become increasingly mature; therefore, it has also been considered in hybrid communication systems. Hybrid THz/RF communication systems and hybrid THz/FSO communication systems have successively appeared.

In 2009, Schober et al. [7] proposed a novel FSO/RF hybrid system architecture to address the problem of unavailable channel state information. It improved the robustness of the hybrid system by applying joint bit-interleaved coded modulation (BICM) to the bit streams transmitted by the RF and FSO subchannels.

In 2009, Nadeem et al. used a GHz-RF link as a backup link and an FSO link to form a hybrid network, considering the high dependence of the FSO link on weather conditions. They investigated the effects of fog, rain, and snow on the FSO/GHz-RF hybrid network to select the GHz RF link with the best complementary characteristics. The simulation results showed that a 40-GHz-RF link significantly improved the availability of the hybrid system link.

In 2010, Abdulhussein et al. [8] proposed an automatic repeat request (ARQ) scheme with rateless coding for an FSO/RF hybrid system. This scheme allowed the efficient utilization of available FSO and RF channel resources, regardless of the changing channel conditions, without requiring any prior rate selection at the transmitter. The simulation results indicated that, under strong atmospheric turbulence conditions, rateless coding was superior to fixed rate coding.

In 2010, Su et al. [9] proposed a switching mechanism based on the network packet-loss rate for an FSO/RF hybrid communication system. It better reflected the current link conditions and enabled the hybrid link to switch efficiently and in real time, thus ensuring uninterrupted data transmissions.

In 2011, Lee et al. [10] studied the performance of an asymmetric dual-hop relay FSO/RF hybrid system in which the RF link followed a Rayleigh distribution and the FSO link followed the gamma–gamma distribution. The relay node used subcarrier intensity modulation to convert the received RF signal into an optical signal. The simulation results showed that with an increase in the signal-to-noise ratio, the outage performance of the hybrid FSO/RF link was infinitely close to that of an RF/RF link.

In 2013, Ansari et al. [11] analyzed an FSO/RF hybrid communication system with pointing errors. They derived the cumulative distribution function, probability density function, and moment generation function of the system and obtained the average bit error rate expression of the hybrid system. Monte Carlo simulation results showed that the bit error performance of the system deteriorated with an increase in the pointing error.

In 2014, Usman et al. [12] studied FSO/RF hybrid communication systems based on hard-switching technology and compared their performances under single and double FSO thresholds. Numerical simulations showed that using FSO/RF hybrid communication effectively improved the system performance, and the dual FSO threshold-switching scheme achieved a performance level similar to that of a single FSO threshold. It also avoided frequent link switching and extended the FSO link life.

In 2015, to mitigate the effects of multipath fading and atmospheric turbulence on an FSO/RF hybrid system, Gao et al. [13] used diversity transmitting at the source and selective combining at the destination, while the relay node was equipped with only one RF receiving antenna and one relay aperture for information forwarding. This provided some theoretical references for the design of an FSO/RF hybrid communication system.

In 2016, Touati et al. [14] investigated the effects of joint fading on hybrid FSO/RF links. The experimental results showed that in the case of strong turbulence or a low signal-to-noise ratio, the performance of the hybrid FSO/RF system was significantly better than that of the FSO alone, and the shortcomings of the FSO link could be effectively overcome.

In 2016, Wang et al. [15] extended multiuser diversity to FSO/RF hybrid links and proposed a point-to-multipoint system composed of hybrid access points, multiple FSO users, and multiple RF users. In each transmission module, the user with the best link quality was selected to communicate with the hybrid access point through the user-scheduling strategy.

In 2017, Hasna et al. [16] proposed an efficient adaptive hybrid FSO/RF communication system, considering actual communication environments, that multiplexed multiple RF users into a single high-speed FSO link, thereby achieving a higher system capacity. However, with a low signal-to-noise ratio, the application of multiuser diversity only slightly improved the system performance.

In 2019, Amirabadi et al. [17] proposed a novel FSO/RF hybrid system model using receive diversity in a parallel single-hop FSO/RF hybrid system. Simulation results showed that the outage probability of the proposed FSO/RF hybrid system was lower than that of the normal FSO and RF systems with receive diversity under different average signal-to-noise ratios. However, the reduction in the power consumption of this FSO/RF hybrid system came at the cost of increased receiver complexity.

In 2019, Esmail et al. [18] experimentally evaluated an all-optical FSO/RF hybrid communication link for 5G applications under severe dust conditions. The

experimental results showed that the RF link had good communication performance in a dense dust environment, whereas the FSO link faced communication interruptions. This proved that the FSO/RF parallel link was a feasible scheme to overcome the dust effect.

In 2019, Yi et al. [19] studied the performance of a dual-hop FSO/RF hybrid communication system based on decoding-and-forwarding in a 5G cloud radio-access network (C-RAN). Users transmitted data to remote RF units through RF links, and the RF units acted as relay systems to forward the received data to the baseband units. Because of small-scale fading and line-of-sight components, the RF link followed a κ-μ shadow distribution.

In 2020, Ashrafzadeh et al. [20] constructed a new hybrid three-hop relay RF/FSO/FSO system to solve the problems of congestion and nonline-of-sight transmissions in urban environments and alleviate the fading effect caused by turbulence. However, the relay nodes had to be equipped with partial or complete FSO transceivers, which significantly increased the system complexity.

In 2020, Zhao et al. [21] studied the performance of an RF/FSO hybrid aeronautical communication system based on selective combining diversity technology. The transmitter of the system did not require a feedback link to obtain the channel state information, which simplified the system design. With a high signal-to-noise ratio, the RF/FSO hybrid parallel communication scheme with receive diversity was better than low-order modulation in improving the system error performance.

In 2021, Khan et al. [22] analyzed an adaptive FSO/RF hybrid system under different weather conditions. Through adaptive modulation and power allocation, the adaptive FSO/RF hybrid system achieved more than a 1-dB performance gain at the minimum power level and adapted to changing weather conditions.

In 2021, Bhardwaj et al. [23] used the THz band for data transmissions between the core network and access point, which proved the feasibility of a THz/RF relay wireless backhaul scheme. Compared to direct transmissions, the spectral efficiency of the scheme was significantly improved by approximately 25%. In the same year, Bhardwaj et al. [24] also studied the performance of a THz/RF relay system with pointing errors under α-μ fading. The diversity order showed that the influence of the pointing error could be reduced by fully adjusting the normalized beam width; however, the attenuation parameters α and μ had greater influence on the system than the normalized beam width.

In 2021, Wu et al. [25], based on the advantages of subcarrier modulation techniques in turbulent atmospheric channels, investigated the performance of an FSO/RF hybrid system with four subcarrier modulation technologies, including differential binary phase-shift keying (DBPSK), noncoherent binary frequency-shift keying (NBFSK), coherent binary phase-shift keying (CBPSK), and coherent binary frequency-shift keying (CBFSK), when pointing errors were considered. Simulation results showed that the hybrid system based on the CBPSK subcarrier modulation technique performed better than the other three subcarrier modulation techniques.

In 2021, Li et al. [26] proposed a hybrid THz/FSO system that realized high-speed and large-capacity data connections between small cells and core networks. The following year, Li et al. [27] studied the performance of a hybrid dual-hop RF/THz system assisted by a reconfigurable intelligent surface (RIS) for the short transmission distance of the THz link. Simulation results showed that the RIS effectively improved the performance of the hybrid RF/THz system.

In 2022, Zafaruddin et al. [28] proposed a hybrid RF/THz communication system with multiple antenna diversity and analyzed its performance using a fixed-gain forwarding protocol with selective combining and a maximal-ratio combining (MRC) diversity receiver; they demonstrated through simulations that increasing the number of antennas improved the system performance.

In 2022, Vishwakarma et al. [29] studied the ergodic capacity of an FSO/RF hybrid system using an adaptive-combining switching scheme. A closed expression of the ergodic capacity under a high signal-to-noise ratio was derived, and the optimal switching threshold and beam width were determined to obtain the optimal ergodic capacity of the adaptive-combining system. The numerical results showed that the ergodic capacity performance of the FSO/RF hybrid system based on adaptive combining was better than those of the MRC and FSO/RF hybrid communication systems based on hard switching.

In 2022, Yang et al. [30] analyzed the performance of a hybrid THz/FSO system. Considering the joint influence of the channel fading and pointing error of the THz and FSO links, the outage probability of a system with impaired hardware was derived and compared with that of the ideal system. The simulation results showed that path loss, channel fading, pointing errors, and hardware impairment led to a significant decrease in the system performance.

In 2023, Singya et al. [31] proposed a backhaul network based on multihop and mesh hybrid THz/FSO technology for high data-rate communication and studied the performance of the system under integrated access and backhaul (IAB) and non-IAB communication settings. The simulation results showed that the IAB mesh network supported multipath routing and significantly improved the system performance.

In 2023, Li et al. [32] studied a multiuser hybrid RF/THz (M-RF/THz) system and derived the performance expressions of the system under fixed-gain amplify-and-forward (AF) and decode-and-forward (DF) protocols. They proved that the diversity order of AF and DF M-RF/THz systems was mainly determined by the pointing error and channel-fading parameters. Increasing the number of users could improve the diversity gain and system performance.

In 2023, Liu et al. [33] proposed a hybrid dual-hop communication system based on cognitive radio with nonorthogonal multiple access (NOMA)-assisted switching, which solved the problems of scarce spectrum resources and low spectral efficiency of communication links. In addition, they studied the influence of the maximum allowable interference power in the main network, the transmission power in the auxiliary transmitter, and the maximum relay transmission power on the link, which provided a new idea for next-generation wireless communication.

System and Channel Models

Channel Model

The hybrid FSO/RF/THz relay system is shown in Fig. 47.1. It is a dual-hop hybrid communication system that provides users with a high speed rate, low latency, and safe and reliable data transmissions. The decode-and-forward (DF) relay mode is adopted. The communication distance from the source node (S) to the relay node (R) is long. With the hybrid FSO/RF communication system based on adaptive combining, the relay node R to the destination node (D) uses a short-distance THz link to access the users. Because FSO only supports line-of-sight communication and RF is used as a backup link for FSO, the transmitters and receivers of both the FSO and RF links are deployed at the top of the building.

For the first-hop S-R hybrid FSO/RF communication system based on adaptive combining [34], when the instantaneous signal-to-noise ratio γ_F of the FSO link receiver is higher than the set threshold γ_{th}, the signal is transmitted through the FSO link. When γ_F of the FSO link drops below γ_{th}, the receiver sends a one-bit feedback message to activate the RF link, and the communication rate of the FSO link is reduced to that of the RF link.

The signal is transmitted simultaneously through the FSO and RF links, and the maximal-ratio combiner is used to combine the signals received from the two links. Therefore, the merging rules of the hybrid FSO/RF communication system based on the adaptive-merging scheme can be expressed as follows:

$$\gamma_c = \begin{cases} \gamma_F, & \gamma_F \geq \gamma_{th} \\ \gamma_F + \gamma_R, & \gamma_F < \gamma_{th} \end{cases}, \tag{47.1}$$

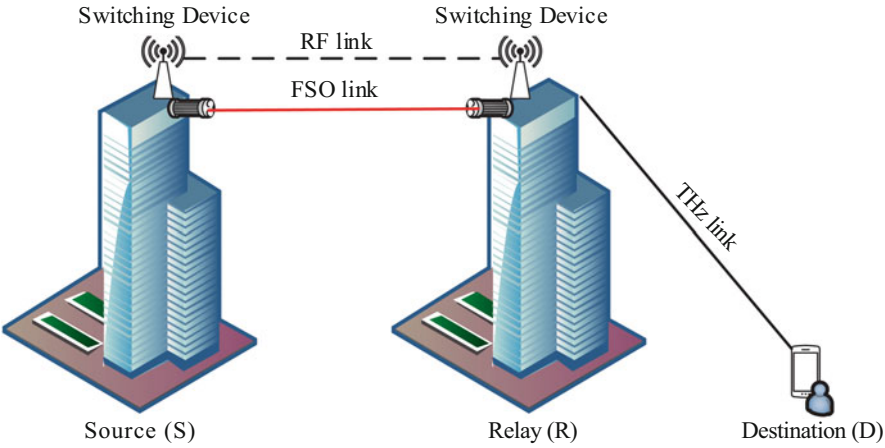


Fig. 47.1 Hybrid FSO/RF/THz relay system model

where γ_c is the output signal-to-noise ratio of the adaptive-combining scheme and γ_R is the instantaneous signal-to-noise ratio of the RF link.

FSO Link

Assuming that the FSO link adopts subcarrier intensity modulation (SIM) and direct detection technology, the optical signal received at relay R can be expressed as

$$y_F = P_F \eta I_l I x + n_F, \tag{47.2}$$

where P_F is the transmission power of the FSO link, η denotes the photoelectric conversion coefficient, and I_l denotes the atmospheric attenuation associated with the FSO link; according to the Beer–Lambert law, $I_l = \exp(-\sigma_l L_F)$, σ_l is the attenuation coefficient, and L_F is the FSO link distance. x represents the signal transmitted at the source node S; n_F represents the additive white Gaussian noise of the FSO link with a mean of 0 and variance of σ_F^2. I represents the effective channel gain of the FSO link, including atmospheric turbulence I_a and pointing errors I_p, where $I = I_a I_p$.

The FSO-link fading I_a caused by turbulence follows a gamma–gamma distribution. Then, the probability density function (PDF) of I_a is

$$f_{I_a}(x) = \frac{2x^{-1}}{\Gamma(\alpha_F)\Gamma(\beta_F)} G_{0\,2}^{2\,0}\left(\alpha_F \beta_F x \middle|_{\alpha_F, \beta_F}^{-}\right), \tag{47.3}$$

where $\Gamma(\cdot)$ represents the gamma function and $G_{p,q}^{m,n}(\cdot)$ denotes the Meijer G-function. The large-scale and small-scale turbulence coefficients, α_F and β_F, are

$$\begin{cases} \alpha_F = \left[\exp\left(\frac{0.49\sigma_{RV}^2}{\left(1+1.11\sigma_{RV}^{12/5}\right)^{7/6}}\right) - 1\right]^{-1} \\ \beta_F = \left[\exp\left(\frac{0.51\sigma_{RV}^2}{\left(1+0.69\sigma_{RV}^{12/5}\right)^{5/6}}\right) - 1\right]^{-1} \end{cases}. \tag{47.4}$$

In Eq. (47.4), the Rytov variance is $\sigma_{RV}^2 = 1.23 C_n^2 k_0^{7/6} L_F^{11/6}$, where the wavenumber $k_0 = \frac{2\pi}{\lambda_F}$ and C_n^2 represents the atmospheric refractive index structure constant.

The fading coefficient is caused by the pointing error during the transmission of the FSO link [35]: $I_p \approx A_{0,F} \exp\left(-\frac{2r^2}{w_{L_{eq}}^2}\right)$, where r represents the radial displacement between the beam center and detector and follows the modified Rayleigh distribution. $A_{0,F}$ represents part of the received optical power at $r = 0$, $A_{0,F} = \mathrm{erf}^2(v)$, $v = \sqrt{\frac{\pi}{2}}\frac{r_a}{w_L}$, r_a denotes the aperture radius, w_L denotes the beam width, and $w_{L_{eq}}$ is the

equivalent beam width, $w_{L_{\mathrm{eq}}}^2 = \frac{w_L^2\sqrt{\pi}\,\mathrm{erf}(v)}{2v\exp(-v^2)}$. Therefore, considering the zero axis pointing error, the PDF of I_{p} is

$$f_{I_{\mathrm{p}}}(x) = \frac{\xi_{\mathrm{F}}^2}{A_{0,\mathrm{F}}^{\xi_{\mathrm{F}}^2}} x^{\xi_{\mathrm{F}}^2-1}, 0 \le x \le A_{0,\mathrm{F}}, \tag{47.5}$$

where the FSO-link pointing error coefficient $\xi_{\mathrm{F}} = \frac{w_{L_{\mathrm{eq}}}}{2\sigma_{j,\mathrm{F}}}$ and $\sigma_{j,\mathrm{F}}$ denotes the standard deviation of the FSO-link jitter.

From $f_I(I) = \int f_{I|I_{\mathrm{a}}}(I|I_{\mathrm{a}}) f_{I_{\mathrm{a}}}(I_{\mathrm{a}})\mathrm{d}I_{\mathrm{a}}$, the PDF of the effective channel gain I of the FSO link under the combined influence of atmospheric turbulence and the pointing error is

$$f_I(I) = \frac{\alpha_{\mathrm{F}}\beta_{\mathrm{F}}\xi_{\mathrm{F}}^2}{A_{0,\mathrm{F}} I_{\mathrm{l}} \Gamma(\alpha_{\mathrm{F}})\Gamma(\beta_{\mathrm{F}})} G_{1,3}^{3,0}\left[\frac{\alpha_{\mathrm{F}}\beta_{\mathrm{F}} I}{A_{0,\mathrm{F}} I_{\mathrm{l}}}\middle|_{\xi_{\mathrm{F}}^2-1,\alpha_{\mathrm{F}}-1,\beta_{\mathrm{F}}-1}^{\xi_{\mathrm{F}}^2}\right]. \tag{47.6}$$

Because the instantaneous and average signal-to-noise ratios of the FSO link are $\gamma_{\mathrm{F}} = \frac{|I|^2 P_{\mathrm{F}}^2 \eta^2}{\sigma_{\mathrm{F}}^2}$ and $\overline{\gamma}_{\mathrm{F}} = \frac{P_{\mathrm{F}}^2\eta^2\xi_{\mathrm{F}}^4}{\sigma_{\mathrm{F}}^2\left(\xi_{\mathrm{F}}^2+1\right)^2} A_{0,\mathrm{F}}^2 I_{\mathrm{l}}^2$, respectively, the PDF of the signal-to-noise ratio γ_{F} of the FSO link can be obtained by using γ_{F} and $\overline{\gamma}_{\mathrm{F}}$ to replace variable I in Eq. (47.6) [14, 36]:

$$f_{\gamma_{\mathrm{F}}}(x) = \frac{\xi_{\mathrm{F}}^2}{2\Gamma(\alpha_{\mathrm{F}})\Gamma(\beta_{\mathrm{F}})x} G_{1,3}^{3,0}\left[\frac{\alpha_{\mathrm{F}}\beta_{\mathrm{F}}\xi_{\mathrm{F}}^2}{\xi_{\mathrm{F}}^2+1}\left(\frac{x}{\overline{\gamma}_{\mathrm{F}}}\right)^{\frac{1}{2}}\middle|_{\xi_{\mathrm{F}}^2,\alpha_{\mathrm{F}},\beta_{\mathrm{F}}}^{\xi_{\mathrm{F}}^2+1}\right]. \tag{47.7}$$

Equation (47.7) is integrated to obtain the cumulative distribution function (CDF) of γ_{F} as follows:

$$F_{\gamma_{\mathrm{F}}}(x) = \frac{2^{\alpha_{\mathrm{F}}+\beta_{\mathrm{F}}-3}\xi_{\mathrm{F}}^2}{\pi\Gamma(\alpha_{\mathrm{F}})\Gamma(\beta_{\mathrm{F}})} G_{3,7}^{6,1}\left(\frac{\alpha_{\mathrm{F}}^2\beta_{\mathrm{F}}^2\xi_{\mathrm{F}}^4 x}{16\left(\xi_{\mathrm{F}}^2+1\right)^2\overline{\gamma}_{\mathrm{F}}}\middle|_{\frac{\xi_{\mathrm{F}}^2}{2},\frac{\xi_{\mathrm{F}}^2+1}{2},\frac{\alpha_{\mathrm{F}}}{2},\frac{\alpha_{\mathrm{F}}+1}{2},\frac{\beta_{\mathrm{F}}}{2},\frac{\beta_{\mathrm{F}}+1}{2},0}^{1,\frac{\xi_{\mathrm{F}}^2+1}{2},\frac{\xi_{\mathrm{F}}^2+2}{2}}\right). \tag{47.8}$$

RF Link

The received signal of the RF link at relay R can be expressed as

$$y_{\mathrm{R}} = \sqrt{P_{\mathrm{R}} h_{\mathrm{l,R}}} h_{\mathrm{R}} x + n_{\mathrm{R}} \tag{47.9}$$

where P_{R} is the transmit power of the RF link, and the path loss of the RF link is $h_{\mathrm{l,R}} = G_{\mathrm{t,R}} + G_{\mathrm{r,R}} - 20\mathrm{lg}\left(\frac{4\pi L_{\mathrm{R}}}{\lambda_{\mathrm{R}}}\right) - \left(\alpha_{\mathrm{oxy}} + \alpha_{\mathrm{rain}}\right)L_{\mathrm{R}}$. $G_{\mathrm{t,R}}$ and $G_{\mathrm{r,R}}$ represent the gain

of the RF-link transmitting and receiving antennas, respectively. L_R and λ_R are the RF-link distance and wavelength, respectively. α_{oxy} and α_{rain} are the attenuation coefficients caused by oxygen and rain absorption, respectively. h_R is the fading coefficient of the RF link, which follows the Nakagami-m distribution, and n_R is the additive Gaussian white noise for the RF link with a mean value of 0 and variance of σ_R^2.

The RF link receives the instantaneous and average signal-to-noise ratios of $\gamma_R = \overline{\gamma}_R |h_R|^2$ and $\overline{\gamma}_R = \frac{P_R h_{l,R}}{\sigma_R^2}$, respectively; therefore, the PDF [37] of γ_R is

$$f_{\gamma_R}(x) = \left(\frac{m}{\overline{\gamma}_R}\right)^m \frac{x^{m-1}}{\Gamma(m)} \exp\left(-\frac{mx}{\overline{\gamma}_R}\right) \tag{47.10}$$

where m denotes the channel fading degree parameter.

Integrating Eq. (47.10) produces the CDF of γ_R as

$$F_{\gamma_R}(x) = \frac{1}{\Gamma(m)} \gamma\left(m, \frac{mx}{\overline{\gamma}_R}\right) \tag{47.11}$$

where $\gamma(\cdot, \cdot)$ denotes the lower bounded incomplete Gamma function.

THz Link

Relay node R uses DF relay forwarding; and then the signal received at destination node D for the second-hop THz link is

$$y_T = \sqrt{P_T} h_{l,T} h_T \widehat{x} + n_T \tag{47.12}$$

where P_T is the transmit power of the THz link, the THz-link path loss $h_{l,T} = \frac{c\sqrt{G_{t,T} G_{r,T}}}{4\pi f_T L_T} \exp\left[-\frac{1}{2}\kappa_a(f_T) L_T\right]$, and c is the speed of light. $G_{t,T}$ and $G_{r,T}$ denote the gains of the transmitting antenna and receiving antenna of the THz link, respectively, and f_T and L_T are the frequency and link distance of the THz link, respectively. $\kappa_a(\cdot)$ denotes the absorption coefficient with respect to temperature, relative humidity, and atmospheric pressure, $\widehat{x}$ denotes the decoded signal at relay R, n_T denotes the additive Gaussian white noise in the THz link with a mean of 0 and a variance of σ_T^2, and h_T denotes the THz-link channel coefficients including the pointing error h_p and the channel fading h_f; $h_T = h_p h_f$.

The THz-link channel coefficients h_T experience α-μ fading with pointing error impairments, so the PDF of h_T, considering the combined effect of channel fading and the pointing error, is

$$f_{h_{\mathrm{T}}}(h_{\mathrm{T}}) = \frac{\xi_{\mathrm{T}}^2 \mu^{\frac{\xi_{\mathrm{T}}^2}{\alpha}} h_{\mathrm{T}}^{\xi_{\mathrm{T}}^2-1}}{A_{0,\mathrm{T}}^{\xi_{\mathrm{T}}^2} \Omega^{\xi_{\mathrm{T}}^2} \Gamma(\mu)} \Gamma\left(\frac{\alpha\mu - \xi_{\mathrm{T}}^2}{\alpha}, \frac{\mu h_{\mathrm{T}}^{\alpha}}{A_{0,\mathrm{T}}^{\alpha} \Omega^{\alpha}}\right) \tag{47.13}$$

where ξ_{T} and $A_{0,\mathrm{T}}$ denote the THz-link pointing error parameter and the partially received optical power, respectively; their solution method refers to the FSO-link pointing error parameter ξ_{F} and the partially received optical power $A_{0,\mathrm{F}}$. α and μ denote the channel-fading parameters, Ω is the α-root mean of the fading channel envelope, and $\Gamma(\cdot, \cdot)$ represents the upper-bound incomplete Gamma function.

The THz link receives the instantaneous signal-to-noise ratio $\gamma_{\mathrm{T}} = \overline{\gamma}_{\mathrm{T}} |h_{\mathrm{T}}|^2$ and the THz-link average signal-to-noise ratio $\overline{\gamma}_{\mathrm{T}} = \frac{P_{\mathrm{T}} h_{\mathrm{l,T}}^2}{\sigma_{\mathrm{T}}^2}$. Therefore, using γ_{T} and $\overline{\gamma}_{\mathrm{T}}$ as substitutes for h_{T} in Eq. (47.13), the PDF [38] of γ_{T} can be obtained as

$$f_{\gamma_{\mathrm{T}}}(x) = \frac{\xi_{\mathrm{T}}^2 \mu^{\frac{\xi_{\mathrm{T}}^2}{\alpha}} x^{\frac{\xi_{\mathrm{T}}^2}{2}-1}}{2A_{0,\mathrm{T}}^{\xi_{\mathrm{T}}^2} \Omega^{\xi_{\mathrm{T}}^2} \Gamma(\mu) \overline{\gamma}_{\mathrm{T}}^{\frac{\xi_{\mathrm{T}}^2}{2}}} \Gamma\left(\frac{\alpha\mu - \xi_{\mathrm{T}}^2}{\alpha}, \frac{\mu x^{\frac{a}{2}}}{A_{0,\mathrm{T}}^{\alpha} \Omega^{\alpha} \overline{\gamma}_{\mathrm{T}}^{\frac{a}{2}}}\right) \tag{47.14}$$

By integrating Eq. (47.14), we obtain the CDF of γ_{T}, as follows:

$$F_{\gamma_{\mathrm{T}}}(x) = \frac{\xi_{\mathrm{T}}^2 \mu^{\frac{\xi_{\mathrm{T}}^2}{\alpha}} x^{\frac{\xi_{\mathrm{T}}^2}{2}}}{\alpha A_{0,\mathrm{T}}^{\xi_{\mathrm{T}}^2} \Omega^{\xi_{\mathrm{T}}^2} \Gamma(\mu) \overline{\gamma}_{\mathrm{T}}^{\frac{\xi_{\mathrm{T}}^2}{2}}} G_{2,3}^{2,1}\left[\frac{\mu x^{\frac{a}{2}}}{A_{0,\mathrm{T}}^{\alpha} \Omega^{\alpha} \overline{\gamma}_{\mathrm{T}}^{\frac{a}{2}}} \middle|_{0, \frac{\alpha\mu - \xi_{\mathrm{T}}^2}{\alpha}, -\frac{\xi_{\mathrm{T}}^2}{\alpha}}^{1-\frac{\xi_{\mathrm{T}}^2}{\alpha}, 1}\right] \tag{47.15}$$

Performance Analysis

Outage Probability

The outage probability refers to the probability that the end-to-end (E2E) signal-to-noise ratio of a communication system will be lower than a set threshold [32, 39]. Therefore, in the first-hop S-R hybrid FSO/RF communication system based on adaptive combining, when the instantaneous signal-to-noise ratio γ_{c} of the output after adaptive combining is lower than the preset outage threshold γ_{out}, the system is in an outage state. The outage probability can be expressed as $P_{\mathrm{out,AC}} = F_{\gamma_{\mathrm{c}}}(\gamma_{\mathrm{out}}) = \Pr[\gamma_{\mathrm{c}} < \gamma_{\mathrm{out}}]$. According to Eq. (47.1), the outage probability of the hybrid FSO/RF communication system based on adaptive combining can be expressed as

$$\begin{aligned} P_{\text{out,AC}} &= F_{\gamma_c}(x) \\ &= \Pr[\gamma_F \geq \gamma_{th}, \gamma_F < x] + \Pr[\gamma_F < \gamma_{th}, \gamma_F + \gamma_R < x] \\ &= \begin{cases} F_1(x), & x \leq \gamma_{th} \\ F_2(x) - F_{\gamma_F}(\gamma_{th}) + F_{\gamma_F}(x), & x > \gamma_{th} \end{cases} \end{aligned} \tag{47.16}$$

where $F_1(x) = \int_0^x f_{\gamma_F+\gamma_R}(t)\mathrm{d}t$ and $F_2(x) = \int_0^{\gamma_{th}} f_{\gamma_F}(t) F_{\gamma_R}(x-t)\mathrm{d}t$. Because the FSO and RF links are statistically independent, $f_{\gamma_F+\gamma_R}(x) = \int_0^x f_{\gamma_F}(t) f_{\gamma_R}(x-t)\mathrm{d}t$. Substituting Eqs. (47.7) and (47.10) into the calculation yields

$$\begin{aligned} f_{\gamma_F+\gamma_R}(x) = \frac{2^{\alpha_F+\beta_F-3}\xi_F^2}{\pi\Gamma(\alpha_F)\Gamma(\beta_F)\Gamma(m)} \left(\frac{m}{\overline{\gamma}_R}\right)^m x^{m-1} \sum_{n=0}^{\infty} (-1)^n \frac{1}{n!} \left(\frac{m}{\overline{\gamma}_R}\right)^n \Gamma(m+n) x^n \\ G_{3,7}^{6,1}\left(\frac{\alpha_F^2 \beta_F^2 \xi_F^4 x}{16\left(\xi_F^2+1\right)^2 \overline{\gamma}_F} \middle| \begin{matrix} 1, \frac{\xi_F^2+1}{2}, \frac{\xi_F^2+2}{2} \\ \frac{\xi_F^2}{2}, \frac{\xi_F^2+1}{2}, \frac{\alpha_F}{2}, \frac{\alpha_F+1}{2}, \frac{\beta_F}{2}, \frac{\beta_F+1}{2}, 1-m-n \end{matrix} \right) \end{aligned} \tag{47.17}$$

Substituting Eq. (47.17) into expressions $F_1(x)$ yield and substituting Eqs. (47.1) and (47.7) into expressions $F_2(x)$ yield

$$\begin{aligned} F_1(x) = \frac{2^{\alpha_F+\beta_F-3}\xi_F^2}{\pi\Gamma(\alpha_F)\Gamma(\beta_F)\Gamma(m)} \left(\frac{m}{\overline{\gamma}_R}\right)^m \sum_{n=0}^{\infty} (-1)^n \frac{1}{n!} \left(\frac{m}{\overline{\gamma}_R}\right)^n \Gamma(m+n) x^{n+m} \\ G_{4,8}^{6,2}\left(\frac{\alpha_F^2 \beta_F^2 \xi_F^4 x}{16\left(\xi_F^2+1\right)^2 \overline{\gamma}_F} \middle| \begin{matrix} 1-n-m, 1, \frac{\xi_F^2+1}{2}, \frac{\xi_F^2+2}{2} \\ \frac{\xi_F^2}{2}, \frac{\xi_F^2+1}{2}, \frac{\alpha_F}{2}, \frac{\alpha_F+1}{2}, \frac{\beta_F}{2}, \frac{\beta_F+1}{2}, 1-m-n, -n-m \end{matrix} \right) \end{aligned} \tag{47.18}$$

$$\begin{aligned} F_2(x) = F_{\gamma_F}(\gamma_{th}) - \frac{2^{\alpha_F+\beta_F-3}\xi_F^2}{\pi\Gamma(\alpha_F)\Gamma(\beta_F)} e^{-\frac{mx}{\overline{\gamma}_R}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{m\gamma_{th}}{\overline{\gamma}_R}\right)^n \sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{mx}{\overline{\gamma}_R}\right)^k \\ \sum_{j=0}^{k} \binom{k}{j} \left(\frac{-\gamma_{th}}{x}\right)^j G_{3,7}^{6,1}\left(\frac{\alpha_F^2 \beta_F^2 \xi_F^4 \gamma_{th}}{16\left(\xi_F^2+1\right)^2 \overline{\gamma}_F} \middle| \begin{matrix} 1-n-j, \frac{\xi_F^2+1}{2}, \frac{\xi_F^2+2}{2} \\ \frac{\xi_F^2}{2}, \frac{\xi_F^2+1}{2}, \frac{\alpha_F}{2}, \frac{\alpha_F+1}{2}, \frac{\beta_F}{2}, \frac{\beta_F+1}{2}, -n-j \end{matrix} \right) \end{aligned} \tag{47.19}$$

Considering the DF relay, the E2E outage probability of the hybrid FSO/RF/THz relay communication system is

$$P_{\text{out}} = 1 - (1 - P_{\text{out,AC}}) \times (1 - P_{\text{out,T}}) \tag{47.20}$$

where $P_{\text{out,T}}$ denotes the outage probability of the second-hop R-D THz link and $P_{\text{out,T}} = F_{\gamma_T}(x)$. Thus, Eq. (47.20) can again be expressed as $P_{\text{out}} = F_{\gamma_c}(x) + F_{\gamma_T}(x) - F_{\gamma_c}(x) \times F_{\gamma_T}(x)$, and the outage probability of the system can be determined by substituting Eqs. (47.15) and (47.16).

Average Bit Error Rate

According to the definition, the average bit error rate (BER) [40] of a communication system can be expressed as $\overline{P}_{\mathrm{e}} = \int_0^{\infty} P(\mathrm{e}|x) f_{\gamma}(x)\mathrm{d}x$, where $P(\mathrm{e}|x)$ denotes the conditional error probability corresponding to different modulation schemes. Therefore, assuming that multiple phase-shift keying (MPSK) modulation is adopted, $P(\mathrm{e}|x) \approx \frac{A}{2}\mathrm{erfc}(\sqrt{x}B)$, and the average BER of the first-hop S-R hybrid FSO/RF communication system based on adaptive combining is

$$\overline{P}_{\mathrm{e,AC}} = \int_0^{\infty} \frac{A}{2}\mathrm{erfc}\left(\sqrt{x}B\right) f_{\gamma_{\mathrm{c}}}(x)\mathrm{d}x \tag{47.21}$$

where $B = \sin(\pi/M)$. When $M = 2$ (BPSK), $A = 1$; when $M > 2$, $A = 2$.

By differentiating the CDF of Eq. (47.16), the PDF of γ_{c} is

$$f_{\gamma_{\mathrm{c}}}(x) = \begin{cases} f_{\gamma_{\mathrm{F}}+\gamma_{\mathrm{R}}}(x), & x \le \gamma_{\mathrm{th}} \\ f_{\gamma_{\mathrm{F}}}(x) + G(x), & x > \gamma_{\mathrm{th}} \end{cases} \tag{47.22}$$

where $G(x) = \int_0^{\gamma_{\mathrm{th}}} f_{\gamma_{\mathrm{F}}}(t) f_{\gamma_{\mathrm{R}}}(x-t)\mathrm{d}t$. Substituting Eqs. (47.7) and (47.10) into the calculation,

$$G(x) = \frac{2^{\alpha_{\mathrm{F}}+\beta_{\mathrm{F}}-3}\xi_{\mathrm{F}}^2}{\pi\Gamma(\alpha_{\mathrm{F}})\Gamma(\beta_{\mathrm{F}})\Gamma(m)}\left(\frac{m}{\overline{\gamma}_{\mathrm{R}}}x\right)^m \sum_{i=0}^{\infty}\frac{1}{i!}\left(-\frac{m}{\overline{\gamma}_{\mathrm{R}}}x\right)^i \sum_{j=0}^{m+i-1}\frac{(-\gamma_{\mathrm{th}})^j}{x^{j+1}} \times \binom{m+i-1}{j} G_{3,7}^{6,1}\left(\frac{\alpha_{\mathrm{F}}^2\beta_{\mathrm{F}}^2\xi_{\mathrm{F}}^4\gamma_{\mathrm{th}}}{16\left(\xi_{\mathrm{F}}^2+1\right)^2\overline{\gamma}_{\mathrm{F}}}\middle|\begin{matrix}1-j, \frac{\xi_{\mathrm{F}}^2+1}{2}, \frac{\xi_{\mathrm{F}}^2+2}{2} \\ \frac{\xi_{\mathrm{F}}^2}{2}, \frac{\xi_{\mathrm{F}}^2+1}{2}, \frac{\alpha_{\mathrm{F}}}{2}, \frac{\alpha_{\mathrm{F}}+1}{2}, \frac{\beta_{\mathrm{F}}}{2}, \frac{\beta_{\mathrm{F}}+1}{2}, -j\end{matrix}\right) \tag{47.23}$$

Substituting Eq. (47.22) into Eq. (47.21) yields

$$\overline{P}_{\mathrm{e,AC}} = \underbrace{\int_0^{\gamma_{\mathrm{th}}} \frac{A}{2}\mathrm{erfc}\left(\sqrt{x}B\right) f_{\gamma_{\mathrm{F}}+\gamma_{\mathrm{R}}}(x)\mathrm{d}x}_{I_1} + \underbrace{\int_{\gamma_{\mathrm{th}}}^{\infty} \frac{A}{2}\mathrm{erfc}\left(\sqrt{x}B\right)\left[f_{\gamma_{\mathrm{F}}}(x) + G(x)\right]\mathrm{d}x}_{I_2} \tag{47.24}$$

Because $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$, the first term I_1 in (47.24) can be expressed as

$$I_1 = \frac{A}{2}\left[\underbrace{\int_0^{\gamma_{\text{th}}} f_{\gamma_{\text{F}}+\gamma_{\text{R}}}(x)\mathrm{d}x}_{F_1(\gamma_{\text{th}})} - \underbrace{\int_0^{\gamma_{\text{th}}} \text{erf}\left(\sqrt{x}B\right) f_{\gamma_{\text{F}}+\gamma_{\text{R}}}(x)\mathrm{d}x}_{I_{11}}\right] \tag{47.25}$$

Substituting Eq. (47.17) into the expression of I_{11} in Eq. (47.25) yields

$$I_{11} = \frac{2^{\alpha_{\text{F}}+\beta_{\text{F}}-2}\xi_{\text{F}}^2}{\pi^{3/2}\Gamma(\alpha_{\text{F}})\Gamma(\beta_{\text{F}})\Gamma(m)}\left(\frac{m\gamma_{\text{th}}}{\overline{\gamma}_{\text{R}}}\right)^m \sum_{n=0}^{\infty}\frac{(-1)^n\left(B\sqrt{\gamma_{\text{th}}}\right)^{2n+1}}{n!(2n+1)}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\left(\frac{m\gamma_{\text{th}}}{\overline{\gamma}_{\text{R}}}\right)^k$$
$$\Gamma(m+k)G_{4,8}^{6,2}\left(\frac{\alpha_{\text{F}}^2\beta_{\text{F}}^2\xi_{\text{F}}^4\gamma_{\text{th}}}{16\left(\xi_{\text{F}}^2+1\right)^2\overline{\gamma}_{\text{F}}}\middle|\begin{matrix}\frac{1}{2}-n-m-k,1,\frac{\xi_{\text{F}}^2+1}{2},\frac{\xi_{\text{F}}^2+2}{2}\\ \frac{\xi_{\text{F}}^2}{2},\frac{\xi_{\text{F}}^2+1}{2},\frac{\alpha_{\text{F}}}{2},\frac{\alpha_{\text{F}}+1}{2},\frac{\beta_{\text{F}}}{2},\frac{\beta_{\text{F}}+1}{2},1-k-m,-n-k-m-\frac{1}{2}\end{matrix}\right) \tag{47.26}$$

Therefore, I_1 can be obtained by substituting Eqs. (47.18) and (47.26) into Eq. (47.25).

Similarly, the second term I_2 in (47.24) can be expressed as

$$\begin{aligned} I_2 = &\underbrace{\int_0^{\infty}\frac{A}{2}\text{erfc}\left(\sqrt{x}B\right)f_{\gamma_{\text{F}}}(x)\mathrm{d}x}_{\overline{P}_{\text{e,F}}} - \frac{A}{2}\underbrace{\int_0^{\gamma_{\text{th}}} f_{\gamma_{\text{F}}}(x)\mathrm{d}x}_{F_{\gamma_{\text{F}}}(\gamma_{\text{th}})} \\ &+\underbrace{\int_0^{\gamma_{\text{th}}}\frac{A}{2}\text{erf}\left(\sqrt{x}B\right)f_{\gamma_{\text{F}}}(x)\mathrm{d}x}_{I_{21}} + \underbrace{\int_{\gamma_{\text{th}}}^{\infty}\frac{A}{2}\text{erfc}\left(\sqrt{x}B\right)G(x)\mathrm{d}x}_{I_{22}} \end{aligned} \tag{47.27}$$

Substituting Eqs. (47.7) and (47.23) into the expressions for $\overline{P}_{\text{e,F}}$, I_{21}, and I_{22} in Eq. (47.27), and obtaining them in turn,

$$\overline{P}_{\text{e,F}} = \frac{2^{\alpha_{\text{F}}+\beta_{\text{F}}-4}A\xi_{\text{F}}^2}{\pi^{3/2}\Gamma(\alpha_{\text{F}})\Gamma(\beta_{\text{F}})}G_{4,7}^{6,2}\left(\frac{\alpha_{\text{F}}^2\beta_{\text{F}}^2\xi_{\text{F}}^4}{16B^2\left(\xi_{\text{F}}^2+1\right)^2\overline{\gamma}_{\text{F}}}\middle|\begin{matrix}1,\frac{1}{2},\frac{\xi_{\text{F}}^2+1}{2},\frac{\xi_{\text{F}}^2+2}{2}\\ \frac{\xi_{\text{F}}^2}{2},\frac{\xi_{\text{F}}^2+1}{2},\frac{\alpha_{\text{F}}}{2},\frac{\alpha_{\text{F}}+1}{2},\frac{\beta_{\text{F}}}{2},\frac{\beta_{\text{F}}+1}{2},0\end{matrix}\right) \tag{47.28}$$

$$I_{21} = \frac{2^{\alpha_{\text{F}}+\beta_{\text{F}}-3}A\xi_{\text{F}}^2}{\pi^{3/2}\Gamma(\alpha_{\text{F}})\Gamma(\beta_{\text{F}})}\sum_{n=0}^{\infty}\frac{(-1)^n\left(B\sqrt{\gamma_{\text{th}}}\right)^{2n+1}}{n!(2n+1)}$$
$$G_{3,7}^{6,1}\left(\frac{\alpha_{\text{F}}^2\beta_{\text{F}}^2\xi_{\text{F}}^4\gamma_{\text{th}}}{16\left(\xi_{\text{F}}^2+1\right)^2\overline{\gamma}_{\text{F}}}\middle|\begin{matrix}\frac{1}{2}-n,\frac{\xi_{\text{F}}^2+1}{2},\frac{\xi_{\text{F}}^2+2}{2}\\ \frac{\xi_{\text{F}}^2}{2},\frac{\xi_{\text{F}}^2+1}{2},\frac{\alpha_{\text{F}}}{2},\frac{\alpha_{\text{F}}+1}{2},\frac{\beta_{\text{F}}}{2},\frac{\beta_{\text{F}}+1}{2},-n-\frac{1}{2}\end{matrix}\right) \tag{47.29}$$

$$I_{22}=\frac{2^{\alpha_F+\beta_F-4}A\xi_F^2}{\pi^{3/2}\Gamma(\alpha_F)\Gamma(\beta_F)\Gamma(m)}\left(\frac{m\gamma_{th}}{\overline{\gamma}_R}\right)^m\sum_{i=0}^{\infty}\frac{1}{i!}\left(\frac{-m\gamma_{th}}{\overline{\gamma}_R}\right)^i\sum_{j=0}^{m+i-1}\binom{m+i-1}{j}(-1)^j$$
$$G_{3,7}^{6,1}\left(\frac{\alpha_F^2\beta_F^2\xi_F^4\gamma_{th}}{16\left(\xi_F^2+1\right)^2\overline{\gamma}_F}\middle|\begin{matrix}1-j,\frac{\xi_F^2+1}{2},\frac{\xi_F^2+2}{2}\\ \frac{\xi_F^2}{2},\frac{\xi_F^2+1}{2},\frac{\alpha_F}{2},\frac{\alpha_F+1}{2},\frac{\beta_F}{2},\frac{\beta_F+1}{2},-j\end{matrix}\right)G_{2,3}^{3,0}\left(B^2\gamma_{th}\middle|\begin{matrix}1,1-m-i+j\\ -m-i+j,0,\frac{1}{2}\end{matrix}\right) \tag{47.30}$$

Therefore, I_2 can be obtained by substituting Eqs. (47.8), (47.28), (47.29), and (47.30) into Eq. (47.27). Then, the derived I_1 and I_2 can be substituted into Eq. (47.24) to obtain the average BER of the first-hop S-R hybrid FSO/RF communication system based on adaptive combining, $\overline{P}_{e,AC}$.

The average BER of the second-hop RD THz link can be expressed as

$$\overline{P}_{e,T}=\int_0^{\infty}\frac{A}{2}\mathrm{erfc}\left(\sqrt{x}B\right)f_{\gamma_T}(x)\mathrm{d}x \tag{47.31}$$

Substituting Eq. (47.14) into Eq. (47.31) yields

$$\overline{P}_{e,T}=\frac{A\xi_T^2\mu^{\frac{\xi_T^2}{\alpha}}}{2\sqrt{\pi}A_{0,T}^{\frac{\xi_T^2}{2}}\Omega^{\xi_T^2}\Gamma(\mu)\overline{\gamma}_T^{\frac{\xi_T^2}{2}}}\frac{2^{\frac{\alpha\mu-\xi_T^2}{\alpha}-\frac{1}{2}}\alpha^{\frac{\xi_T^2}{2}-1}}{(2\pi)^{\frac{\alpha}{2}}B^{\frac{\xi_T^2}{2}}}$$
$$\times G_{2+2\alpha,4+\alpha}^{4,2\alpha}\left[\frac{\mu^2\alpha^{\alpha}}{4B^{\alpha}A_{0,T}^{2\alpha}\Omega^{2\alpha}\overline{\gamma}_T^{\alpha}}\middle|\begin{matrix}\frac{1-\xi_T^2/2}{\alpha},\ldots,\frac{\alpha-\xi_T^2/2}{\alpha},\frac{1-\xi_T^2/2-1/2}{\alpha},\ldots,\frac{\alpha-\xi_T^2/2-1/2}{\alpha},\frac{1}{2},1\\ 0,\frac{1}{2},\frac{\alpha\mu-\xi_T^2}{2\alpha},\frac{\alpha\mu-\xi_T^2+1}{2\alpha},\frac{1-\xi_T^2/2-1}{\alpha},\ldots,\frac{\alpha-\xi_T^2/2-1}{\alpha}\end{matrix}\right] \tag{47.32}$$

Considering the DF relay, the E2E average BER of the hybrid FSO/RF/THz relay communication system is

$$\overline{P}_e=\overline{P}_{e,AC}+\overline{P}_{e,T}-2\times\overline{P}_{e,AC}\times\overline{P}_{e,T} \tag{47.33}$$

Thus, the average BER of the system can be determined by substituting Eqs. (47.24) and (47.32) into Eq. (47.33).

Simulation Results

To analyze the performance of the hybrid FSO/RF/THz relay system and verify the accuracy of the analysis results, simulations were performed based on the outage probability and BER expressions derived in section "Performance Analysis," and the system simulation parameter settings are shown in Table 47.1.

Figures 47.2, 47.3, and 47.7 show the outage probability and bit error performance of the first-hop S-R hybrid FSO/RF system in the hybrid FSO/RF/THz relay

Table 47.1 Simulation parameter settings

FSO link		RF link		THz link	
Parameter	Value	Parameter	Value	Parameter	Value
Wavelength λ_F	1550 nm	Carrier frequency f_R	28 GHz	Carrier frequency f_T	300 GHz
Opt-to-elect. conversion η	1	Link distance L_R	500 m	Link distance L_T	100 m
Strong turbulence C_n^2	1×10^{-12} $m^{-2/3}$	Transmit antenna gain $G_{t,R}$	44 dBi	Transmit antenna gain $G_{t,T}$	55 dBi
Moderate turbulence C_n^2	1×10^{-13} $m^{-2/3}$	Receive antenna gain $G_{r,R}$	44 dBi	Receive antenna gain $G_{r,T}$	55 dBi
Turbulence coefficient (strong) α_F, β_F	4.7206, 1.2066	Oxygen absorption α_{oxy}	15.1 dB/km	Humidity ϕ	50%
Turbulence coefficient (moderate) α_F, β_F	5.6137, 4.0062	Rain attenuation α_{rain}	0 dB/km	Atmosphere pressure p	101,325 Pa
Link distance L_F	500 m			Temperature T	296 K
Receiving aperture r_a	20 cm			Jitter standard deviation $\sigma_{j,T}$	5–15 cm
Jitter standard deviation $\sigma_{j,F}$	5–20 cm				

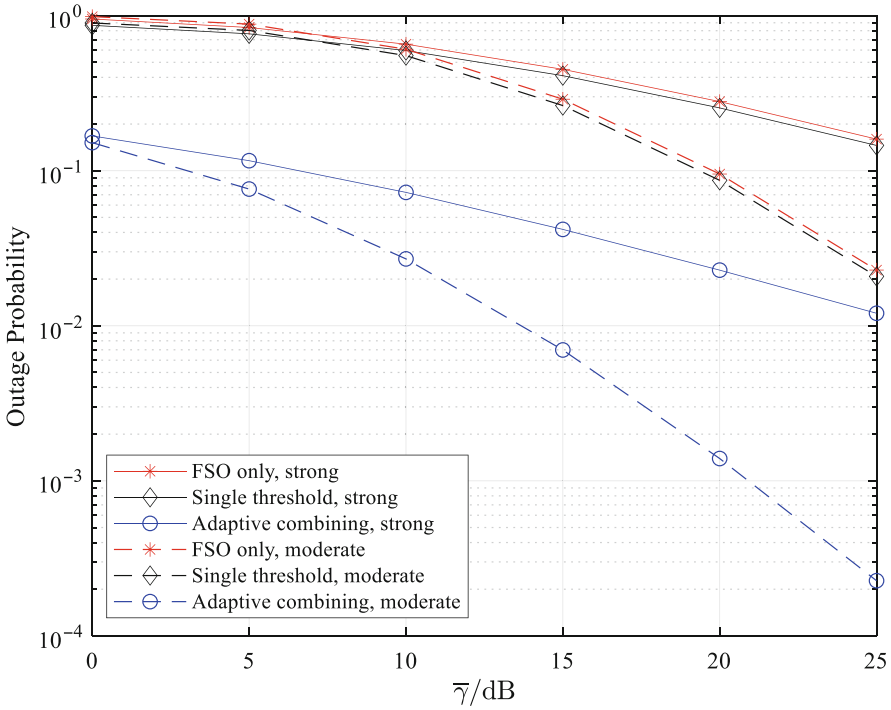


Fig. 47.2 Comparison of the outage performance of the adaptive-combining scheme, single-threshold switching scheme, and single-link FSO

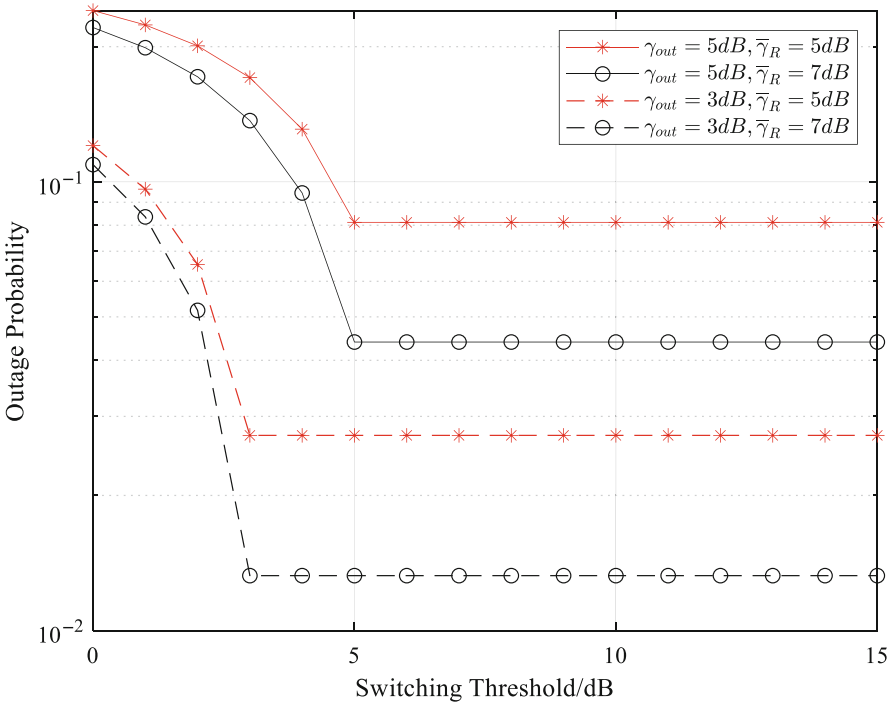


Fig. 47.3 Relationship between the outage probability and switching threshold of the adaptive-combining scheme

system, where $\overline{\gamma}_F = \overline{\gamma}$ is the average signal-to-noise ratio for the FSO link. Figures 47.4, 47.5, 47.6, 47.8, and 47.9 show the performance of the hybrid FSO/RF/THz relay system under the influence of different parameters, where the average signal-to-noise ratios of the FSO, RF, and THz links are $\overline{\gamma}_F = \overline{\gamma}_R = \overline{\gamma}_T = \overline{\gamma}$, respectively.

Figure 47.2 compares the outage performances of the single-link FSO, single-threshold switching, and adaptive-combining schemes under moderate and strong turbulence conditions, where $m = 1$, $\overline{\gamma}_R = 8$dB, $\xi_F = 4.5747$, $\gamma_{th} = 10$dB, and $\gamma_{out} = 3$dB. As shown in Fig. 47.2, compared with moderate turbulence, the outage probability of the system under strong turbulence conditions increases, indicating that an increase in the turbulence intensity leads to the deterioration of the system outage performance.

In addition, it can be observed from Fig. 47.2 that the outage performance of the adaptive-combining system is better than those of the single-link FSO and single-threshold switching scheme. The main reason for this phenomenon is that when the FSO-link quality deteriorates, the single-link FSO has no backup link. The FSO link must be used for transmission all the time, whereas the single-threshold switching scheme can activate the RF link for transmission at this time. However, link switching may occur frequently, which reduces the system transmission

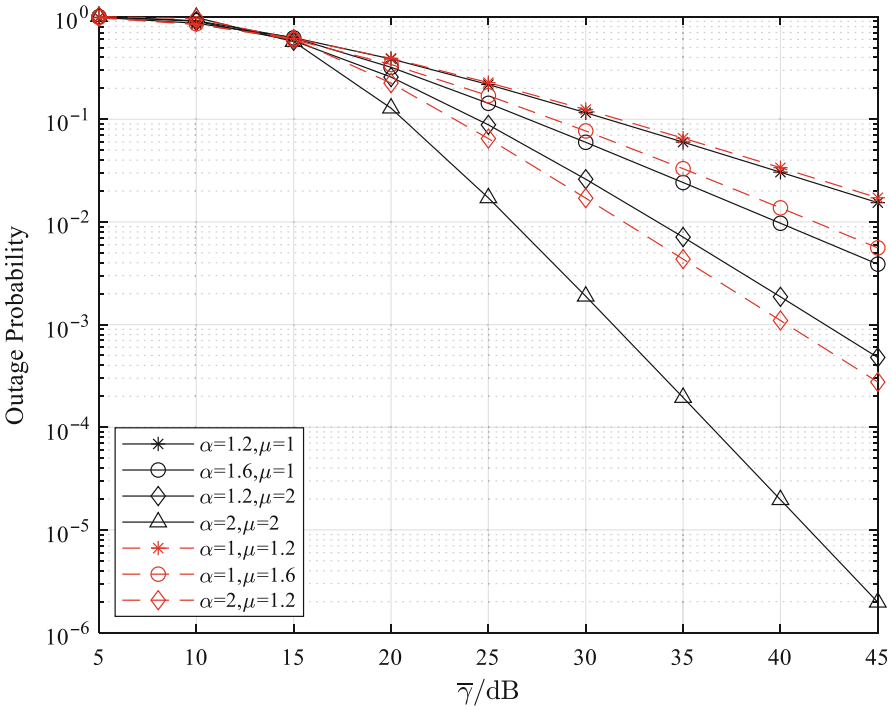


Fig. 47.4 Outage probability of the system under the influence of different fading parameters

performance. The adaptive-combining scheme overcomes the shortcomings of both, so it has better system performance.

Figure 47.3 shows the relationship curve between the outage probability and the switching threshold of the adaptive-combining scheme, where $C_n^2 = 1 \times 10^{-12}\mathrm{m}^{-2/3}$ (moderate turbulence), $m = 2$, $\overline{\gamma}_\mathrm{F} = 10\mathrm{dB}$, and $\xi_\mathrm{F} = 4.5747$. As shown in Fig. 47.3, when $\gamma_\mathrm{out} > \gamma_\mathrm{th}$, the system outage probability gradually decreases with an increase in the switching threshold. When $\gamma_\mathrm{out} \leq \gamma_\mathrm{th}$, the system outage probability reached its minimum value, and the outage probability remains constant. Therefore, the switching threshold must be set above the outage threshold to obtain the best outage performance.

Moreover, increasing the average signal-to-noise ratio $\overline{\gamma}_\mathrm{R}$ of the RF link can improve the system outage performance. This is because as $\overline{\gamma}_\mathrm{R}$ increases, the RF-link quality increases, thereby providing a better-quality backup link for the FSO link.

Figure 47.4 shows the relationship between the average signal-to-noise ratio and the hybrid FSO/RF/THz outage probability under different fading parameters, where $C_n^2 = 1 \times 10^{-12}\mathrm{m}^{-2/3}$ (moderate turbulence), $\xi_\mathrm{F} = 4.5747$, $\xi_\mathrm{T} = 3.4450$, $m = 1$, and $\gamma_\mathrm{out} = 5\mathrm{dB}$. It can be seen from Fig. 47.4 that as the values of α and μ increase, the fading severity of the THz link decreases, improving the outage performance of the system.

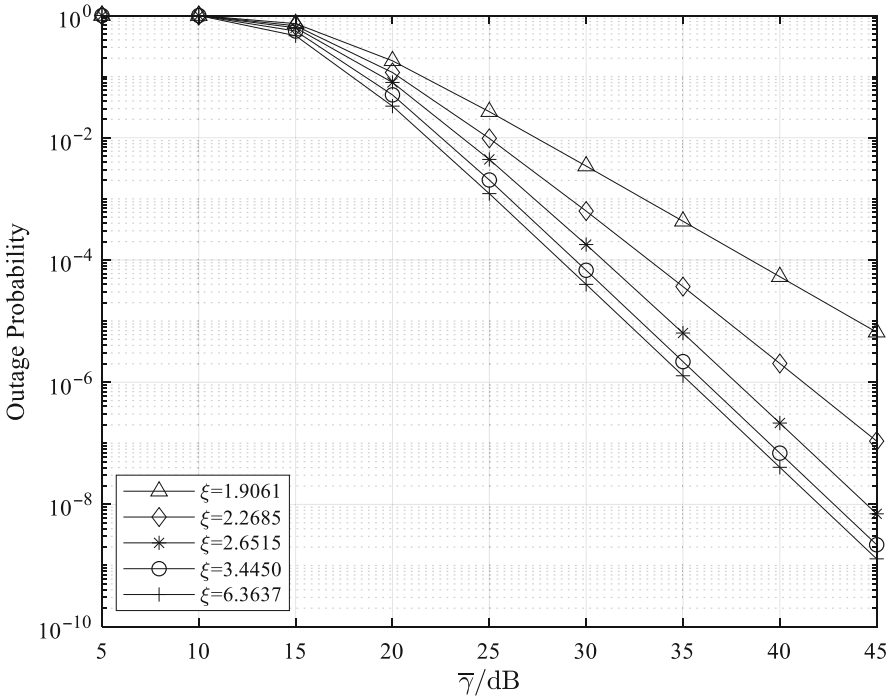


Fig. 47.5 Outage probability of the system under different pointing errors

In addition, the outage performance of the system is compared when varying α, and then when varying μ: $\{\alpha = 1.2, \mu = 1\}$, $\{\alpha = 1.6, \mu = 1\}$, $\{\alpha = 2, \mu = 1\}$ and $\{\alpha = 1, \mu = 1.2\}$, $\{\alpha = 1, \mu = 1.6\}$, $\{\alpha = 1.2, \mu = 2\}$. It can be observed that when the degree of fading is comparable, the outage probability of the system is lower when $\alpha > \mu$ than when $\alpha < \mu$.

Figure 47.5 shows the curves of the average signal-to-noise ratio versus the outage probability of the hybrid FSO/RF/THz system under different pointing errors, where $\xi_F = \xi_T = \xi$, $C_n^2 = 1 \times 10^{-12}\text{m}^{-2/3}$ (moderate turbulence), $m = 1$, $\alpha = 3$, $\mu = 2$, and $\gamma_{out} = 5\text{dB}$. From Fig. 47.5, it can be seen that the smaller the value of pointing error ξ, the greater the impact on the system, resulting in a higher probability of system interruption; therefore, increasing the pointing error can improve the outage performance of the system. However, after the pointing error has increased to a certain value, continuing to increase its value will have a reduced effect on the system outage performance.

Figure 47.6 shows the relationship between the outage probability and average signal-to-noise ratio. The outage performances of a single FSO, single RF, single THz, first-hop adaptive-combining hybrid FSO/RF system, and end-to-end hybrid FSO/RF/THz relay system were studied, where $\xi_F = 4.5747$, $\xi_T = 3.4450$, and $\gamma_{out} = 5\text{dB}$.

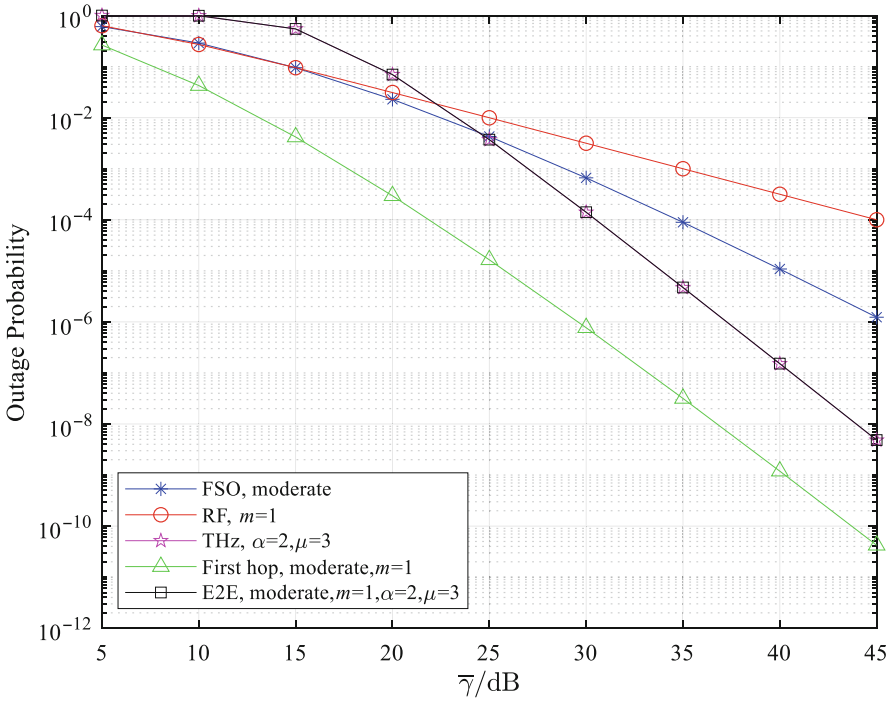


Fig. 47.6 Relationship between the system outage probability and average signal-to-noise ratio

It can be observed from Fig. 47.6 that the outage performance of the THz link is the worst in the low signal-to-noise ratio region. With an increase in the average signal-to-noise ratio, the quality of the THz link continuously improves. In the high signal-to-noise ratio region, the outage performance of the THz link is better than that of the FSO and RF links. In addition, it can be seen from Eq. (47.20) that the outage probability of the dual-hop relay system is determined by the hop with the worst outage performance. Therefore, the outage performance of the end-to-end hybrid FSO/RF/THz relay system shown in Fig. 47.6 is consistent with that of the second-hop THz link.

Figure 47.7 compares the BER performance of the system with a single FSO link, single threshold-switching scheme, and adaptive-combining scheme under moderate and strong turbulence conditions, where $M = 2$, $m = 1$, $\overline{\gamma}_{\mathrm{R}} = 5\mathrm{dB}$, $\xi_{\mathrm{F}} = 4.5747$, and $\gamma_{\mathrm{th}} = 7\mathrm{dB}$.

Figure 47.7 shows that reducing the turbulence intensity of the FSO link effectively improves the error performance of the system. It is also observed that the error performance of the adaptive-combining system is better than those of the single-link FSO and single-threshold switching schemes under both moderate and strong turbulence conditions.

Figure 47.8 shows the relationship curve between the average signal-to-noise ratio and BER of the hybrid FSO/RF/THz system under different fading parameters

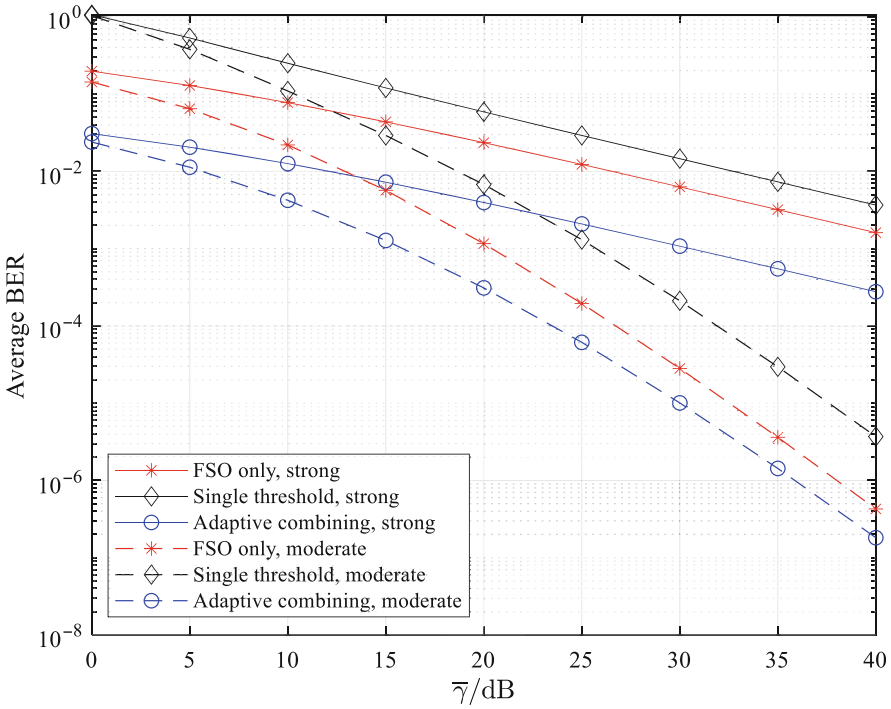


Fig. 47.7 Comparison of the BER performance of the adaptive-combining scheme with those of the single-threshold switching scheme and single-link FSO

and pointing errors, where $\xi_F = \xi_T = \xi$, $C_n^2 = 1 \times 10^{-12} m^{-2/3}$ (moderate turbulence), $M = 2$, $m = 1$, and $\gamma_{th} = 7$dB.

From Fig. 47.8, it can be seen that the pointing-error parameters of the FSO and THz links, as well as the fading parameters α and μ of the THz link, have an impact on the BER performance of the system. The smaller the values of α, μ, or ξ, the more serious the fading and pointing errors of the system. Increasing the value of α, μ, or ξ can reduce the BER of the system and improve the communication quality of the system.

Figure 47.9 shows the relationship between the average BER and average signal-to-noise ratio of the system. The BER performances of a single FSO, single RF, single THz, first-hop adaptive-combining hybrid FSO/RF system, and end-to-end hybrid FSO/RF/THz relay system were studied, where $M = 2$, $\xi_F = 4.5747$, $\xi_T = 3.4450$, and $\gamma_{th} = 7$dB.

As can be seen in Fig. 47.9, in the low signal-to-noise ratio region, the BER performance of the end-to-end hybrid FSO/RF/THz relay system remains consistent with that of the THz link until it enters the high signal-to-noise ratio region, where it approaches that of the first-hop adaptive-combining hybrid FSO/RF system until the two are basically overlapped. This verifies the accuracy of Eq. (47.33).

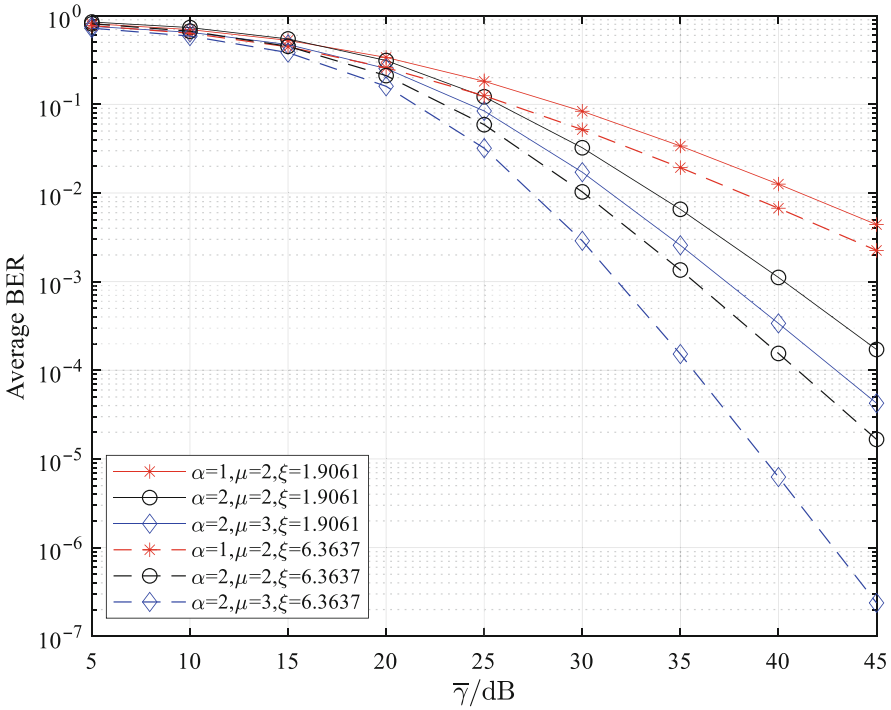


Fig. 47.8 Average BER of the system under different fading parameters and pointing errors

Conclusion

This study designed a hybrid FSO/RF/THz relay communication system that used hybrid FSO/RF communication based on adaptive combining as the backhaul network before the relay node. It then connected to users through the THz link, which provided safe, reliable, low-latency, and high-speed data transmissions.

Considering atmospheric turbulence and pointing errors in the FSO link, and fading and pointing errors in the THz link, the performance equations of the hybrid FSO/RF/THz communication system were derived using the obtained statistical characteristics of the FSO, RF, and THz links.

The simulation results showed that compared with the single-threshold switching scheme and the single-link FSO system, the adaptive-combining scheme in the first hop of the relay system provided better system performance, and the optimal switching threshold of the adaptive-combining scheme was obtained. Moreover, by evaluating the performance of the hybrid FSO/RF/THz relay system under different fading conditions and pointing errors, it was found that both channel fading and pointing errors reduced the system performance. In the high signal-to-noise ratio region, the hybrid FSO/RF/THz relay system provided better outage and BER performance.

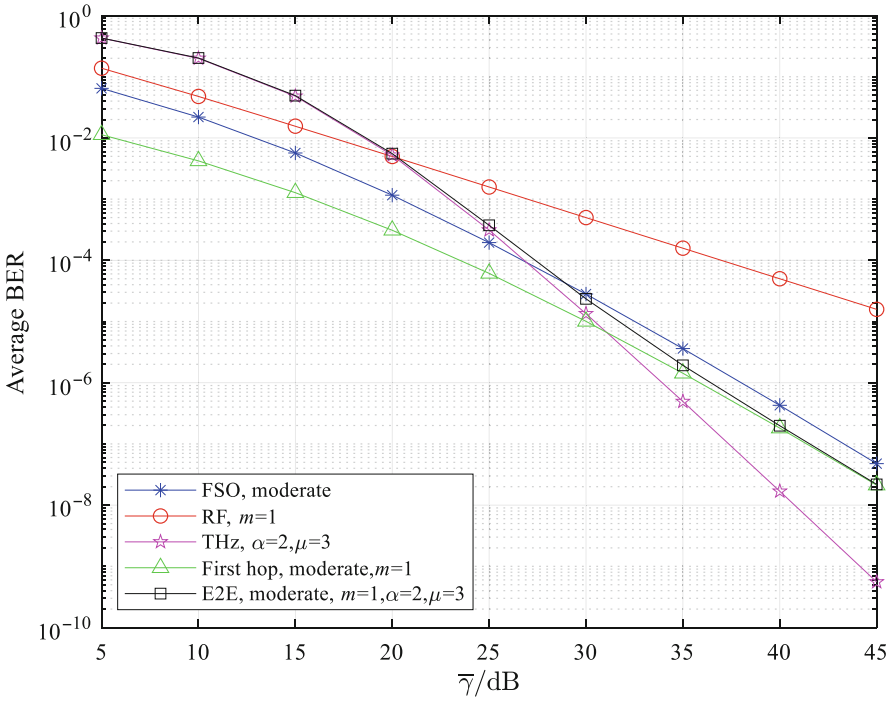


Fig. 47.9 Relationship between the average BER and the average signal-to-noise ratio of the system

References

1. Samy R, Yang H C, Rakia T, et al. Space-air-ground FSO networks for high-throughput satellite communications [J]. IEEE Communications Magazine, 2022, 60(12): 82–87.
2. Zedini E, Soury H, Alouini M S. On the performance analysis of dual-hop mixed FSO/RF systems [J]. IEEE Transactions on Wireless Communications, 2016, 15(5): 3679–3689.
3. Fawaz W, Abou R C, Assi C. UAV-aided cooperation for FSO communication systems [J]. IEEE Communications Magazine, 2018, 56(1): 70–75.
4. Nadeem F, Kvicera V, Awan M S, et al. Weather effects on hybrid FSO/RF communication link [J]. IEEE journal on selected areas in communications, 2009, 27(9): 1687–1697.
5. Bhardwaj P, Zafaruddin S M. On the performance of multihop THz wireless system over mixed channel fading with shadowing and antenna misalignment [J]. IEEE Transactions on Communications, 2022, 70(11): 7748–7763.
6. Zhu M, Zhang J, Hua B, et al. Ultra-wideband fiber-THz-fiber seamless integration communication system toward 6G: architecture, key techniques, and testbed implementation [J]. Science China Information Sciences, 2023, 66(1):1–18.
7. He B, Schober R. Bit-interleaved coded modulation for hybrid RF/FSO systems [J]. IEEE Transactions on Communications, 2009, 57(12): 3753–3763.
8. Abdulhussein A, Oka A, Nguyen T, et al. Rateless coding for hybrid free-space optical and radio-frequency communication [J]. IEEE Transactions on Wireless Communications, 2010, 9(3): 907–913.

9. Su J, Chen H. Research on switching technology of hybrid FSO/RF system based on network packet loss rate [J]. Optical communication technology, 2011, 35(10):28–31.
10. Lee E, Park J, Han D, et al. Performance analysis of the asymmetric dual-hop relay transmission with mixed RF/FSO links [J]. IEEE Photonics Technology Letters, 2011, 23(21): 1642–1644.
11. Ansari I S, Yilmaz F, Alouini M S. Impact of pointing errors on the performance of mixed RF/FSO dual-hop transmission systems [J]. IEEE Wireless Communications Letters, 2013, 2(3): 351–354.
12. Usman M, Yang H C, Alouini M S. Practical switching-based hybrid FSO/RF transmission and its performance analysis [J]. IEEE Photonics Journal, 2014, 6(5): 1–13.
13. Yang L, Hasna M O, Gao X. Performance of mixed RF/FSO with variable gain over generalized atmospheric turbulence channels [J]. IEEE Journal on Selected Areas in Communications, 2015, 33(9): 1913–1924.
14. Touati A, Abdaoui A, Touati F, et al. On the effects of combined atmospheric fading and misalignment on the hybrid FSO/RF transmission [J]. Journal of Optical Communications and Networking, 2016, 8(10): 715–725.
15. Chen L, Wang W, Zhang C. Multiuser diversity over parallel and hybrid FSO/RF links and its performance analysis [J]. IEEE Photonics Journal, 2016, 8(3): 1–9.
16. Yang L, Hasna M O, Ansari I S. Unified performance analysis for multiuser mixed η-μ and M-distribution dual-hop RF/FSO systems [J]. IEEE Transactions on Communications, 2017, 65(8): 3601–3613.
17. Amirabadi M A, Vakili V T. A novel hybrid FSO/RF communication system with receive diversity [J]. Optik, 2019, 184(3): 293–298.
18. Esmail M A, Ragheb A M, Fathallah H A, et al. 5G-28 GHz signal transmission over hybrid all-optical FSO/RF link in dusty weather conditions [J].IEEE Access, 2019, 7(2):24404–24410.
19. Yi X, Shen C, Yue P, et al. Performance of decode-and-forward mixed RF/FSO system over κ−μ shadowed and exponentiated Weibull fading [J]. Optics Communications, 2019, 439(1): 103–111.
20. Ashrafzadeh B, Soleimani-Nasab E, Zaimbashi A, et al. Outage performance of mixed RF-FSO systems over DGG and Nakagami-m channels [J]. IEEE Wireless Communications Letters, 2020, 9(12): 2135–2139.
21. Liu W, Zhao S, Wang X, et al. Performance analysis of a hybrid FSO/RF aeronautical communication system based on SC diversity [J]. Laser and Infrared, 2020, 50(05):525–531.
22. Khan M N, Kashif H, Rafay A. Performance and optimization of hybrid FSO/RF communication system in varying weather [J]. Photonic Network Communications, 2021, 41(1): 47–56.
23. Bhardwaj P, Zafaruddin S M. Performance of dual-hop relaying for THz-RF wireless link [C]. IEEE Vehicular Technology Conference, Helsinki, 2021: 1–5.
24. Bhardwaj P, Zafaruddin S M. Performance of Dual-Hop Relaying for THz-RF Wireless Link Over Asymmetrical α-μ Fading [J]. IEEE Transactions on Vehicular Technology, 2021, 70(10): 10031–10047.
25. Wu Y, Mei H, Dai C, et al. Performance analysis of hybrid free space optical/radio frequency communication systems [J]. Laser & Optoelectronics progress, 2021, 58(17):203–209.
26. Li S, Yang L, Hu Y. Performance analysis of mixed THz/FSO relaying systems [C]. International Conference on Wireless Communications and Signal Processing (WCSP), Changsha, 2021: 1–5.
27. Li S, Yang L, Cui Q, et al. Performance analysis of RIS-assisted mixed RF/THz systems [J]. Journal of Communications, 2022, 43(01):49–58.
28. Bhardwaj P, Zafaruddin S M. Performance of Hybrid THz and Multiantenna RF System With Diversity Combining [J]. IEEE Systems Journal, 2022.
29. Vishwakarma N, Swaminathan R. On the capacity performance of hybrid FSO/RF system with adaptive combining over generalized distributions [J]. IEEE Photonics Journal, 2022, 14(1): 1–12.

30. Li S, Yang L, Zhang J, et al. Mixed THz/FSO relaying systems: Statistical analysis and performance evaluation [J]. IEEE Transactions on Wireless Communications, 2022, 21(12): 10996–11010.
31. Singya P K, Makki B, Errico D A, et al. High-rate Reliable Communication using Multi-hop and Mesh THz/FSO Networks [J]. arXiv preprint arXiv:2304.01643, 2023.
32. Li S, Yang Y, Yang L, et al. On the Study of Multiuser Mixed Dual-Hop RF/THz Systems [J]. IEEE Transactions on Vehicular Technology, 2023, 72(7):9175–9188.
33. Liu R, Wang Z, Wang X, et al. Performance Analysis of Soft-Switching FSO/THz-RF Dual-Hop AF-NOMA Link Based on Cognitive Radio [J]. Photonics, 2023, 10(10): 1086–1099.
34. Shah S, Siddharth M, Vishwakarma N, et al. Adaptive-combining-based hybrid FSO/RF satellite communication with and without HAPS [J]. IEEE Access, 2021, 9: 81492–81511.
35. Boluda-Ruiz R, García-Zambrana A, Castillo-Vázquez C, et al. Novel approximation of misalignment fading modeled by Beckmann distribution on free-space optical links [J]. Optics express, 2016, 24(20): 22635–22649.
36. Djordjevic G T, Petkovic M I, Cvetkovic A M, et al. Mixed RF/FSO relaying with outdated channel state information [J]. IEEE Journal on selected areas in Communications, 2015, 33(9): 1935–1948.
37. Sharma S, Madhukumar A S, Swaminathan R. Effect of pointing errors on the performance of hybrid FSO/RF networks [J]. IEEE Access, 2019, 7: 131418–131434.
38. Li S, Yang L. Performance analysis of dual-hop THz transmission systems over α-μ fading channels with pointing errors [J]. IEEE Internet of Things Journal, 2021, 9(14): 11772–11783.
39. Rakia T, Yang H C, Alouini M S, et al. Outage analysis of practical FSO/RF hybrid system with adaptive combining [J]. IEEE Communications Letters, 2015, 19(8): 1366–1369.
40. Singya P K, Makki B, D'Errico A, et al. Hybrid FSO/THz-based backhaul network for mmWave terrestrial communication [J]. IEEE Transactions on Wireless Communications, 2022, 22(7): 4342–4359.

Orbital Angular Momentum Beam Techniques

48

Contents

Introduction 1756
Basic OAM Concepts 1756
OAM Properties 1757
OAM Beam Types 1757
Research Progress on Vortex Beam Generation 1758
Single Vortex Beam Generation 1758
Generation of Superimposed Vortex Beams 1760
Research Progress on Vortex Beam Transmissions 1763
Vortex Beam Transmission Characteristics 1763
Beam Transmission Characteristics of Partially Coherent Vortex Beams 1765
Vector Vortex Beam Transmission Characteristics 1766
Adaptive Wavefront Correction for Vortex Beams 1767
Vortex Beam Separation Measurement 1770
OAM Technology in Optical Wireless Communication Systems at Xi'an University of Technology 1774
Generation of Vortex Beams Carrying OAM 1778
Vortex Beam OAM Transmission 1786
Vortex Beam Separation and Detection 1802
Vortex Beam OAM Applications 1810
Research Outlook of OAM Technology for Optical Wireless Communication Systems 1817
References 1818

Abstract

The transmission capacity of optical wireless communication systems has approached the Shannon limit. The emergence of orbital angular momentum (OAM), as a new lateral space-dimension resource, can exponentially improve the capacity and spectral efficiency of optical wireless communication networks. Adopting OAM-related technologies can provide a potential solution for realizing ultra-high-speed and high-capacity cross-scene optical communications in the future. After comparing and summarizing the relevant domestic and foreign OAM research results, this chapter introduces the OAM research work in the

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_48

field of optical wireless communication at Xi'an University of Technology. It is mainly divided into space generation, propagation characteristics, separation and detection, and OAM beam applications. Finally, the future development trends and prospects of OAM technology in the field of optical wireless communication are discussed, providing new ideas and reference values for subsequent researchers to explore in this field.

Keywords

Optical wireless communication · Orbital angular momentum · Beam generation · Propagation characteristics · Separation detection

Introduction

Allen's significant discovery of a spiral-structured light-beam distribution carrying orbital angular momentum (OAM) at Woerdman's Optical Quantum Laboratory in 1992 [1] triggered a wave of OAM research. This unique spatial-domain distribution has promising applications for expanding the channel capacity. The specially structured optical field carrying OAM has also been widely used in scientific fields, such as ultra-high-capacity optical communication [2], particle capture [3], optical metrology [4], and quantum science [5].

Optical wireless communication uses a laser as a carrier wave to transmit data, such as video and voice. It has the advantages of a high transmission rate, flexible networking, strong resistance to electromagnetic interference, etc. It also has good applications in solving the "last mile" bottleneck problem and complex and harsh environment communication [6].

When a vortex beam is transmitted in free space, its absorption, scattering, and attenuation by atmospheric turbulence can cause beam spreading, wandering, and scintillation, resulting in a large communication bit error rate and even communication interruptions [7]. Therefore, the study of OAM technology with a high transmission rate, high security, and strong immunity is one of the research hotspots in the field of optical wireless communication.

Basic OAM Concepts

It is well known that light, an electromagnetic wave, has a wave–particle duality. Its fluctuation is familiar to us, while its particle nature, including energy and momentum, is less understood. Light has linear momentum in the longitudinal direction along the propagation direction, as well as spin angular momentum (SAM) associated with polarization properties, and transverse-spin angular momentum perpendicular to the propagation direction, which exists independently of the SAM.

OAM Properties

Beams with OAM are called vortex beams, and they have been widely studied and explored by the academic community in recent years. Figure 48.1 shows the spiral-wavefront distribution of a vortex beam with different topological charges. The uncertainty of the phase of the spiral wavefront causes coherent-phase extinction of the light intensity. This forms a phase singularity at the center of the vortex beam, which is the point where the light intensity is zero; the OAM is also zero at this time. The optical field of the vortex beam can be expressed as $E(x, y) = u(r, z) \exp(il\theta) \exp(-ikz)$, where $u(r, z)$ is the amplitude function, k is the wavenumber, and z is the beam propagation distance.

Each vortex beam has a spiral phase factor term $\exp(il\theta)$, where θ is the spatial azimuthal angle; l is the topological charge, that is, one cycle around the phase singularity, and the change in phase is l times 2π, producing l winding spiral wavefronts. Each photon in the beam carries $l\hbar$ OAM ($\hbar = h/2\pi$ is the reduced Planck constant); that is, the optical OAM corresponds to an infinite number of eigenmodes, which can construct Hilbert spaces of infinite dimensions; thus, the information capacity and security can be theoretically increased significantly [8].

OAM Beam Types

Common vortex beams are generally classified into three types: Laguerre–Gaussian (LG), Bessel–Gaussian (BG), and Airy vortex (AV). All these beams carry OAM, but they have different characteristics. For example, the BG beam has

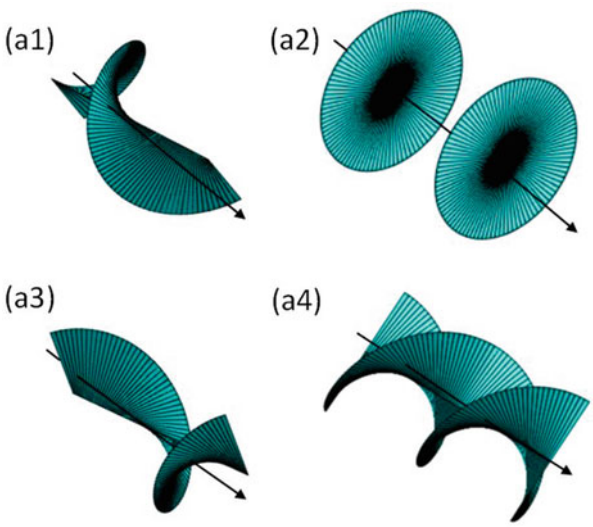


Fig. 48.1 Helical wavefront distribution with different topological charges. (a1) $l = -1$; (a2) $l = 0$; (a3) $l = 1$; and (a4) $l = 2$

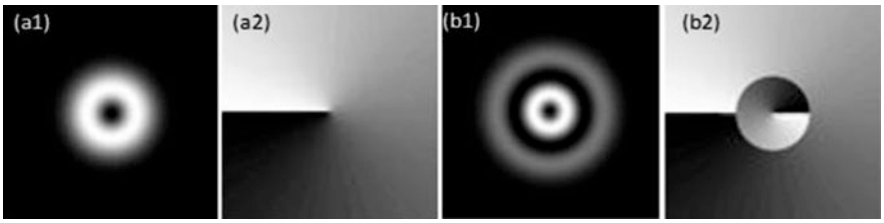


Fig. 48.2 Intensity and phase distribution of LG beams under different l and p. (a1) LG_0^1 beam intensity distribution; (a2) LG_0^1 phase distribution; (b1) LG_1^1 beam intensity distribution; and (b2) LG_1^1 phase distribution

non-diffraction characteristics and good directionality, and the AV beam has self-acceleration and self-bending characteristics.

The LG beam has the characteristics of a Gaussian beam, is easy to prepare in the laboratory, and is one of the most commonly used vortex beam sources. Different combinations of the topological charge l and radial index p can form different forms of LG_{lp} beams. Figure 48.2 shows the light intensity and phase distribution of LG beams at different l and p. The beam profile has a doughnut-shaped distribution, and the number of light rings is equal to $p + 1$.

Research Progress on Vortex Beam Generation

Single Vortex Beam Generation

In 1994, Barnett et al. explored the relationship between the axis symmetric beam and OAM under the non-paraxial approximation. They found that the beam carried OAM as long as it contained the $\exp(il\theta)$ term through theoretical derivation. This laid the foundation for introducing various OAM beams in addition to the Laguerre–Gaussian beam [9].

In 1989, Coullet et al. used a laser resonant cavity to directly generate OAM beams; however, this method required high axisymmetry of the resonant cavity and large intracavity losses, which made it difficult to generate high-quality vortex beams [10].

In 2013, Lee et al. demonstrated direct LG mode generation at the first Stokes wavelength of a diode-side-pumped Nd:$GdVO_4$ self-Raman laser, which was the first intracavity frequency conversion using self-Raman scattering (SRS) [11]. In 2014, Lee et al. demonstrated that the emission of a Stokes wavelength doubled into a 586-nm vortex wave using intracavity Nd:$GdVO_4$ crystals generated from SRS, with a final emission efficiency of 727 mW and a diode-to-visible light conversion efficiency of 4% [12].

In 1992, Heckenberg et al. generated OAM vortex beams using computer-generated holograms (CGH). Holographic gratings can also be generated by methods such as etching scratches on glass plates or holographic photography [13].

In 2012, Escuti et al. designed a forked polarization grating (FPG) that could efficiently generate OAM beams and convert between modes; the relative power between modes could also be controlled by polarization. This method had a diffraction efficiency higher than 90% and was applicable to any wavelength except ultraviolet and infrared [14].

In 2019, Rosales et al. used multilevel OAM holograms to generate structured OAM light fields, illustrating theoretically and experimentally that at least three layers of spiral phase plate (SPP) phases were required to generate OAM beams, whereas the combination with phase diagrams could delay their retention time in three dimensions [15].

In 1993, Beijersbergen et al. used a mode converter consisting of two cylindrical lenses to convert a Hermite–Gaussian (HG) beam into a high-purity LG beam [16].

In 2002, Shvedov et al. used a single optical wedge to produce a single-mode OAM vortex beam. Producing multiple wedges stacked according to certain rules produced higher-order modes [17].

In 2016, Forbes et al. combined spatial light modulators (SLMs) for OAM beam generation and used SLM modulation to generate different types and forms of OAM beams by modifying and writing digital holograms in real time [18].

In 2012, Koh et al. generated a series of convergent OAM electron beams with topological charges up to 90 using a spiral zone plate (SZP) with micromachining technology [19].

In 2021, Hu et al. used a digital micromirror device (DMD) to generate and characterize complex vector modes. They generated vector beams, such as LG, Ince–Gaussian, Mathieu–Gaussian, and parabolic–Gaussian, using a binary encoding scheme based on random spatial multiplexing. Finally, vector-mode characterization techniques, such as the real-time Stokes polarization reconstruction method, were described [20].

In 2015, Padgett et al. designed a generalized photon sieve to generate OAM beams. It reduced the number of pinholes from thousands to hundreds, compared to the original photon sieve, and could generate LG beams with different radial and angular indices, as well as superposition states [21].

In 2017, Devlin et al. used a two-dimensional metasurface consisting of dynamic and geometric phases to generate OAM beams. The metasurface was a J-plate composed of rectangular nanoantennas of different sizes, which allowed any orthogonal polarization state to be converted to independent OAM modes, compared to previous Q-plates, and could be used with high-power laser beams [22].

In 2021, Dorrah et al. designed a total angular momentum (TAM) board that enabled two main classes of functions. The first type was a polarization-switchable device that switched between two generated vortex beams by changing the incident orthogonal polarization state. The other type was effective for any incident polarization state and allowed flexible control of the optical OAM, SAM, phase, polarization, and amplitude along the propagation direction [23].

In 2019, Sheng et al. designed a flat-plate spiral phase plate (SPP) to generate vortex beams; it could flexibly adjust the OAM angular quantum number and SPP

height of the generated vortex beams. The accuracy of the generated beams increased with an increase in the phase order [24].

In 2014, Brunet et al. added a material with a refractive index lower than that of the cladding in the annular region of a ring fiber, and simulated the generation and transmission of 28 OAM modes, while the experiment accommodated 36 modes [25]. In 2017, Pidishety et al. used an all-fiber mode-selective coupler consisting of a single-mode fiber and a ring-core fiber to excite OAM beams by direct phase-matched coupling with a final excitation-mode purity of 75% [26].

In 2020, Hassan et al. proposed a circular-core photonic crystal fiber (PCF) with a large air hole in the center, which supported stable transmissions of up to 38 OAM modes with a larger bandwidth and flat dispersion variation [27]. In the same year, Israk et al. designed a large-bandwidth PCF consisting of three twisted and coiled structural cavities, which supported up to 56 OAM modes with flat dispersion variation and limited losses below 1.7407×10^{-11}dB/m [28].

In 2020, Cognée et al. demonstrated that a plasmonic photon-resonant microdisk cavity using an addressed waveguide could produce a correlated beam with controlled OAM and SAM [29]. The following year, Butt et al. proposed and compared three μ-ring resonant-cavity structures for standard strip, slot, and hybrid plasma waveguides. They found that silicon-strip waveguide ring resonators were simple to fabricate but had low sensitivity; hybrid plasma had the best sensitivity; and the slot waveguide was superior in terms of quality factor and detection limit [30].

Generation of Superimposed Vortex Beams

In 2003, Stabinis et al. superimposed two intersecting LG beams interferometrically, resulting in a more pattern-rich composite optical vortex. The number and position of the composite vortices could be changed by varying the propagation length of the vortices, topological charge of each beam, phase difference, intersection angle, and amplitude ratio [31].

In 2006, Galvez et al. superimposed two LG beams coaxially to produce a composite vortex beam with a novel phase singularity. They focused on the effect of the optical intensity ratio of the two LG beams on the light intensity distribution of the superimposed state of the LG beams [32].

In 2009, Baumann et al. studied the effect of the relative amplitude and phase on the off-axis propagation of the colinear superposition of LG beams and provided an explanation using the Gouy phase [33].

In 2009, Yang et al. studied the colinear composite superposition of planar and annular vortices and proposed a method to obtain the location and distribution properties of such colinear superposition vortices [34].

In 2010, Ando et al. analyzed the structural properties of different phase singularities of double and multiple LG-mode superposition states with different angular and radial indices, and verified the singularity properties by theoretical derivations and experiments [35].

In 2012, Fang et al. generated fractional-order double-ring vortex beams theoretically and experimentally, and double-ring beams with different spacings and notches were obtained by varying the topological charge number. The results showed that the fractional-order double-ring vortex beam could be controlled individually between them; it had more control diversity than the integer-order and single fractional-order beams [36].

In 2015, Zhu et al. proposed a pattern search-assisted iterative (PSI) algorithm using a generated phase diagram. It simultaneously generated multi-OAM mode beams with up to 100 randomly spaced OAM modes and controlled their relative power distributions with a diffraction efficiency higher than 93%, along with a low standard deviation and relative root-mean-square error [37].

In 2016, Huang et al. used an out-of-cavity mode-selection method to achieve the coaxial superposition of multiple LG beams using a modified CGH to produce a high-quality composite vortex beam [38].

In 2020, Mateusz et al. proposed a method for generating high-symmetry composite vortex beams. The method allowed the intensity relations between the structures to be controlled so that they could rotate, while being able to superimpose more than two such structures. Their results showed that the optical vortex superposition of two such structures generated a highly symmetric composite vortex beam [39].

In 2021, Kumar et al. designed a hybrid binary-forked grating (hBFG), which was computationally generated by removing the central region around the corresponding forked dislocation and replacing it with a BFG corresponding to different topological charges. Optimizing the geometric parameters of this grating allowed the generation of a composite vortex beam consisting of three different topological charges and higher-order radial indices, while controlling the intensity and radius of the individual modes [40].

In 2022, Kotlyar et al. studied the superposition of several identical parallel single-ring LG beams and found that when the weighting factor of the superposition was real, the net topological charge of the superposition was equal to the sum of the individual components, regardless of whether it was for the initial plane or the far field. The phase delay between the beams changed the number of superimposed net topological charges [41] (Table 48.1).

Optical OAM has been studied since 1992. In the beginning, OAM beams were generated in a laser resonant cavity [10], that is, an optical vortex laser. This active scheme is generally divided into direct-generation [42], ring-beam pumping [43, 44], central damaged-cavity [45, 46], and thermally induced membrane-aperture methods [47, 48].

A new method of generating computer-controlled vortex lasers [49, 50] has emerged in recent years. It applies SLMs to act as a piece of resonant mirror inside the laser cavity, while the computer enables the flexible customization of the vortex beam with arbitrary properties. These methods of actively generating vortex beams from solid-state laser cavities require special components designed inside the cavity or higher pump optical power. They are simple and easy to operate; however, they cannot generate vortex beams with simultaneous high conversion efficiency, high orders, and multimodality.

Table 48.1 Domestic and foreign research progress on vortex beam generation

Year	Researcher	Research progress
1989	Coullet	Generated a beam in a laser resonator
1992	Heckenberg	CGH
1993	Beijersbergen	Mode converter
2002	Shvedov	Single light wedge
2003	Stabinis	Superposition of interference
2006	Galvez	Effect of light intensity ratio on LG beam superposition state
2009	Baumann	Effect of phase difference on composite vortex beams
2009	Yang	Interference superposition of planar vortex beams
2010	Ando	LG beam superposition state with different topological charges and radial indices
2012	Fang	Superposition of fractional double vortex beams
2012	Koh	SZP
2012	Escuti	Fork-shaped polarization grating
2013	Lee	Diode side-pumped self-Raman Nd:$GdVO_4$ laser
2014	Brunet	Hollow-core step fiber
2014	Brunet	Hollow-core step fiber
2015	Zhu	PSI algorithm generates the phase diagram
2015	Padgett	Generalized photon sieve
2016	Huang	Improved computer hologram to generate multi-LG coaxial superposition
2016	Forbes	SLM with modified CGH in real time
2017	Devlin	Two-dimensional metasurface
2019	Sheng	Flat-plate SPP
2020	Hassan	Circular-core PCF
2020	Cognée	Plasma-resonant microdisk cavity for waveguide addressing
2020	Mateusz	Highly symmetrical composite vortex beams
2021	Hu	DMD represents complex vector patterns
2021	Dorrah	TAM plate
2021	Kumar	Improved hBFG generation
2022	Kotlyar	Parallel single-ring LG superposition

A passive scheme, which mainly uses various optical components outside the resonant cavity to convert other types of beams into vortex beams, is implemented using various and flexible methods, such as CGH [13–15], SZP [19], mode converters [16], optical wedges [17], SPP [24], SLM [18, 53], DMD [20], and photonic sieves [21]. Among these, the CGH diffraction method is easy to operate and inexpensive; however, its diffraction efficiency is low. Methods to optimize the CGH quality using iterative algorithms have also been proposed, such as the Gerchberg–Saxton (GS) iterative algorithm [51] and the pixelwise hybrid input-output (PHIO) method [52].

The SZP method is similar to CGH, but the diffraction uses spherical waves instead of plane waves and is more suitable for the field of electron microscopy. Mode converters produce beams with high purity; however, their operation is

complicated. The optical-wedge block system is simple and reliable, but the components require high processing accuracy. An SPP generates high beam purity but can only produce OAM beams with fixed modes and parameters.

A DMD is inexpensive and can produce OAM modes at high switching rates, but its diffraction efficiency is not ideal. Photon sieves have a wider available waveband and a higher spatial resolution; however, the diffraction efficiency is not high. SLM is convenient for adjusting parameters and has a higher diffraction efficiency; however, the energy threshold is limited.

In the q-plate and J-plate [22], a birefringent crystal allows the transition from the local polarization state to the OAM mode, and generates a higher-purity OAM beam. The J-plate can be used with a high-power laser beam. Finally, metamaterials, metasurfaces [22, 23], fiber-optic devices [25–28], and photonic integrated devices [29, 30] have evolved. Such systems are compact, easy to integrate, and can generate a variety of efficient OAM modes.

In terms of superposition-state vortex beams, the classical method of an SLM-loaded CGH is still mainly used [36–38, 40]. Only the forms of the superposition states vary, from the initial double-superposition state of petal-like and circular hole splitting to the superposition of three, four, or even more LG beams. The types of superposition include fractional-order superposition as well as radial-index superposition.

These multispecies and multimodal superimposed vortex beams can also be controlled by the beam parameters of a single state to transform the superimposed light field structure and light intensity ratio [31–33, 39, 41]. This injects new vitality and lays a solid foundation for enhancing the transmission capacity of large-scale optical wireless OAM communication systems.

Research Progress on Vortex Beam Transmissions

Vortex Beam Transmission Characteristics

In 2001, the International Telecommunication Union (ITU) proposed an ITU-Radiocommunication Sector (ITU-R) atmospheric structure-constant model capable of varying with height to study vortex beam transmissions in turbulence with a slanted path [54].

In 2005, Paterson et al. studied the detection-probability distribution of a vortex light field after turbulent scattering and found that the narrower the beam, the smaller the OAM scattering, and the wider the beam, the smaller the scintillation [55].

In 2008, Anguita et al. used the Kolmogorov atmospheric turbulence spectrum as a model to analyze the system capacity of multiplexed OAM systems when transmitted in free space and proposed a bit error rate (BER) evaluation method [56].

In 2008, Gbur et al. investigated the transmission of vortex beams in weak-to-strong turbulence and showed that the topological charge of the beam was, in principle, a robust quantity that was not lost after long-distance transmission and could be used as a carrier for optical communication [57].

In 2009, Tyler et al. studied the disturbances of OAM beams after turbulent atmospheric transmissions, where the quantitative derivation of the quantum information loss rate was encoded on single-photon OAM states in turbulent transmissions [58].

In 2009, Pu et al. theoretically and experimentally demonstrated the beam expansion of vortex beams in atmospheric turbulence and found that vortex beams were less affected by turbulence than non-vortex beams [59].

In 2011, Pu et al. also investigated the transmission characteristics of elliptical vortex beams at different turbulence intensities. It was found that the smaller the ratio of the short axis to the long axis of the ellipse, the smaller the scintillation index; this indicated that the choice of elliptical vortex beams for transmission could reduce the scintillation index [60]. In 2012, the group investigated the far-field diffraction intensity distribution of off-axis LG beams using the Fresnel diffraction theory and determined the law of the topological charge sign affecting the direction of the singularity movement [61].

In 2012, Lukin et al. numerically simulated the beam spreading of low-order vortex beams in atmospheric turbulence and found that the average spreading of vortex beams was smaller than that of Gaussian beams. In addition, the larger the topological charge number, the smaller the beam spreading and the less it was affected by turbulence [62].

In 2013, Aksenov et al. investigated the wander variance of LG beams in turbulent flows using numerical simulations. The results showed that the spot-wander variance of the LG beam was smaller when the topological charge number was larger, the radial index was larger, the outer scale was smaller, and the beam wander of the Gaussian beam was larger than that of an LG beam [63].

In 2014, Banakh et al. studied the LG beam spreading effect in atmospheric turbulence and simulated the intensity distribution of the LG beam after propagation. The results showed that the turbulent spreading of the LG beam relative to the diffraction spreading did not depend on the mode but on the diffraction region at the emission aperture; instead, the absolute turbulent spreading of the LG beam increased with increasing topological charge [64].

In 2016, Ge et al. studied the relative beam width variation of LG beams in turbulence and found that a longer transmission distance, stronger turbulence, smaller topological charge, larger beam waist, and shorter wavelength led to a larger beam width variation and LG beam spreading when influenced by turbulence [65].

In 2016, Filimonov et al. numerically simulated and analyzed the OAM variation when an LG beam propagated in atmospheric turbulence and studied the effect of beam and receiver axis alignment errors on the mean and variance of OAM fluctuations [66].

In 2017, Porfirev et al. studied the propagation characteristics of vortex beams in a random aerosol medium. It was found that, at short distances, vortex beams were mostly less stable than Gaussian beams; however, at long distances, vortex beams were more stable, probably because of their ability to recover after passing an obstacle [67].

In 2018, Tian et al. conducted a numerical simulation of a longitudinal OAM multiplexing (LOAMM) system under atmospheric turbulence, using a stepwise Fourier-transform method. The results showed that the detection probability spectrum of the system was inversely proportional to the OAM modes, turbulence intensity, and propagation distance under finite-aperture conditions, and the performance of the LOAMM system was better than that of OAM space-division multiplexing (SDM) [68].

In 2019, Luo et al. used the generalized Huygens–Fresnel principle and the Rytov approximation to derive expressions for the transmission intensity of arbitrary-order vortex beams and their arrays in turbulence. They numerically simulated the beam spreading and evolution to obtain the mean-square beam widths for different beam and turbulence parameter variations. They concluded that vortex beam arrays were less affected by turbulence than single-mode vortex beams, and that radial vortex beam arrays eventually evolved into Gaussian-like beams [69].

In 2020, Yan et al. analyzed the far-field optical-intensity distributions of super-Gaussian and Gaussian vortex beams in turbulence using the Fresnel diffraction principle and obtained the variation in the distortion-dispersion degree of the two beams with different parameters [70].

In 2022, Lazrek et al. simulated the evolution of the beam spreading and mean intensity of a vortex cosine-hyperbolic Gaussian beam (vChGB) in ocean turbulence, based on the extended Huygens–Fresnel diffraction principle, and analyzed the effect of parameters, such as the mean-square temperature and dissipation rate of the temperature to the salinity-fluctuation ratio, on the beam propagation [71].

Beam Transmission Characteristics of Partially Coherent Vortex Beams

In 2009, Cai et al. derived an analytical formula for the cross-spectral density of partially coherent standard and elegant LG beams, after passing through a paraxial-axis ABCD optical system, and compared their light intensity and spreading characteristics in free space; they were found to be closely related to the initial coherent state of the source [72, 73].

In 2011, Cui et al. analyzed the angular width and beam transmission factor (M2 factor) of a partially coherent LG beam after transmission in atmospheric turbulence, using the extended Huygens–Fresnel principle and second-order moments of the Wigner distribution function (WDF) [74].

In 2015, Tang et al. investigated the effects of isotropic and anisotropic random media on the transmission characteristics of MGSMV beams in terms of the light intensity distribution and spatial-coherence distribution [75].

In 2017, Yao et al. discussed the effect of non-Kolmogorov turbulence on the beam width and angular spreading of multi-Gaussian–Schell-model vortex (MGSMV) beams and found that the anti-turbulence performance of MGSMV beams was better than that of multi-Gaussian–Schell-model (MGSM) and Gaussian–Schell-model (GSM) beams [76].

In 2018, Gbur et al. introduced the twisted-vortex Gaussian–Schell mode beam (tvGSM), a new partially coherent OAM vortex beam determined by the potential vortex order of the random beam and the "bending" parameter of the beam ensemble. The transmission characteristics and OAM density of the beam were also analyzed [77].

In 2021, Hyde et al. proposed a twisted space-time optical vortex (STOV) beam, derived the mutual-coherence function and linear and angular momentum densities, and simulated the synthesis of a STOV beam and its propagation in free space with different degrees of coherence [78].

In 2022, Hricha et al. studied the focusing characteristics and focal shift of a partially coherent vortex cosine-hyperbolic Gaussian beam (PCvChGB) when it passed through a lens system and concluded that the spatial-coherence length, Fresnel number, and beam parameters had a significant effect on the mean intensity distribution and focal shift of the beam-focusing region [79].

Vector Vortex Beam Transmission Characteristics

In 2009, Haus et al. studied the light intensity distribution model, polarization degree, and scintillation index of vector vortex beams after transmission through atmospheric turbulence, in combination with random phase-screen theory. They compared the transmission characteristics with those of standard Gaussian beams and scalar vortex beams and found that the scintillation of vector vortex beams was less than that of scalar vortex beams under the same light source conditions [80].

In 2013, Martinez-Herrero and Prado discussed the evolution of the polarization structure of partially coherent radially polarized vortex beams [81].

In 2015, Jiang et al. studied the polarization characteristics of elliptically polarized vortex beams in turbulent flows and analyzed the effects of the polarization ellipticity, topological charge number, and beam width on the average polarization intensity and degree [82]. In the same year, Cheng et al. introduced a linearly polarized butterfly–Gaussian vortex beam with a butterfly–Gaussian structure in the far-field plane and discussed the effects of parameters, such as topological charge number, linear polarization angle, off-axis distance, and scattering length, on the far-field diffraction structure of a linearly polarized butterfly–Gaussian vortex beam [83].

In 2016, Guo et al. derived the cross-spectral density matrix of a partially coherent radially polarized vortex (PCRPV) beam after an ABCD optical system and found that the vortex phase reduced the degradation of the intensity distribution caused by coherence, as well as depolarization. They also experimentally generated a PCRP vortex beam [84].

In 2016, Eyyuboğlu et al. studied the scintillation properties of vector vortex beams in a strong turbulence region by using a random phase-screen method. It was found that the scintillation effect gradually diminished with an increasing topological charge number, but the rate of diminution gradually slowed [85].

In 2018, Tang et al. introduced a new radially polarized beam, the multiple cosine Gaussian–Schell-model (MCSGSM) beam, and also analyzed the light intensity, coherence degree, polarization degree, and polarization state of the MCSGSM beam in non-Kolmogorov turbulence. They found that its polarization state had a self-splitting property [86].

In 2020, Joshi et al. theoretically and experimentally investigated the coherent polarization variation of a high-order vector vortex beam. The results showed that the transverse degree of polarization (DOP) of a partially coherent vector vortex beam (PCVVB) depended on the spatial correlation of the beam, whereas the state of polarization (SOP) remained constant [87].

In 2021, Joshi et al. studied the statistical properties of PCVVBs embedded in V points and found that the modulation of the spatial coherence length in the source plane could effectively control the coherence intensity, as well as the DOP and SOP distributions [88].

In 2022, Manisha et al. studied the far-field intensity, SOP, and DOP distributions of a non-isotropic PCVVB, both theoretically and experimentally, and found that the variations in these distributions were related to the Poincaré–Hopf index and non-isotropic parameters of the input beam [89]. In the same year, Arora et al. theoretically and experimentally demonstrated the perturbation of a uniform linearly polarized beam to a vector singularity. The amount of radial displacement of the beam center of mass caused by this perturbation could indicate the strength of the perturbation [90].

Adaptive Wavefront Correction for Vortex Beams

In 2009, Starikov et al. used a Shack–Hartmann (SH) wavefront sensor with a singular reconstruction technique and a bimorphic piezoelectric-ceramic reflector to correct the phase of distorted vortex beams, which greatly improved the Strehl ratio and optical-system resolution [91]. In 2012, Garanin et al. used a flexible deformation mirror and the SH algorithm for wavefront correction of continuous- and pulsed-wave lasers. The results showed a significant improvement in the corrected Strehl ratio and an order-of-magnitude reduction in the beam aberration [92].

In 2012, Leach et al. used two aberration-correction methods, the SH algorithm and the Gerchberg–Saxton (GS) algorithm, to correct a distorted OAM beam after turbulence and found that both methods improved the purity of the single OAM state and the free-space optical (FSO) channel capacity, but the GS algorithm was somewhat better [93].

In 2007, Jesacher et al. used the GS algorithm and a spatial light modulator loaded with its hologram to correct an LG beam, which was better for low-order mode correction [94].

In 2018, Baranek et al. corrected a vortex beam optical aberration by introducing an optimized spiral phase modulation with the GS algorithm and combining it with the SLM into an adaptive correction system. The results showed that the accuracy

and efficiency of the GS-algorithm aberration correction could be significantly improved by using the vortex image spot as the target-intensity mode during the iterative process [95].

In 2015, Xie et al. proposed the stochastic parallel-gradient descent (SPGD) algorithm, based on a Zernike polynomial, to correct multiple distorted OAM beams. The crosstalk between modes was reduced by 5 dB, and the Zernike polynomial coefficients of the corrected modes were obtained from the feedback loop [96].

In 2018, Xu et al. investigated a dual-correction method using a combination of the liquid-crystal adaptive optics (AO) technique and phase-diversity (PD) algorithm. The AO correction technique compensated for the distortion caused by turbulent perturbations for the first correction. Then, the PD algorithm was used to correct the residual wavefront distortion [97].

In 2020, Gao et al. designed a turbulence-aberration correction convolutional neural network (TACCNN) model, which was able to correct and compensate for the turbulence aberration of a vector vortex beam by supervised learning of a large number of experimental samples. The mode purity was improved from 19% to 70% [98].

In 2021, Zhan compared the correction effect of the SH, SPGD, and GS algorithms through an improved underwater-turbulence model using experimental studies and found that the GS algorithm had the best correction effect, the SPGD algorithm the second best, and the SH algorithm the worst [99] (Table 48.2).

The propagation properties of OAM beams in atmospheric turbulence have been gradually studied and explored by a wide range of scholars since 2000. At the beginning, turbulence effects were studied, such as beam spreading [58, 59, 62, 64, 65, 69, 71], beam wander [58, 61, 63, 111], light intensity scintillation distribution [55, 58, 60], energy attenuation, and spiral spectrum dispersion [58, 100]. Coherent optics has also developed rapidly, and vortex beams have been extended from fully coherent to partially coherent [72–79], scalar to vector light [80–90], and integer order to fractional order.

Studies on the transmission properties of different light sources, such as elliptical vortices [60, 109], off-axis vortices [61], flat-top vortices [101], standard and elegant vortices [102], array vortices [69], super-Gaussian [70], vChGB [71, 103], and others [104, 105, 108, 109], have also appeared. The evaluation indices also range from the optical-intensity phase distribution, spiral-spectral distribution, relative beam width, and scintillation index to the transmission factor, complex coherence, beam quality, polarization characteristics, mode probability density, and scattering characteristics [112]. However, most of these properties are still at the theoretical and simulation stage, and further exploration and research in experimental studies are needed [106, 107].

Adaptive wavefront correction gradually developed from the SH algorithm [91–93], with wavefront sensors at the beginning, to the GS algorithm [94, 95], SPGD algorithm [96, 110], and PD algorithm [97], without wavefront sensors. The evaluation indexes also ranged from the optical-intensity correlation coefficient,

Table 48.2 Domestic and foreign research progress on the propagation characteristics of vortex beams

Year	Person/ organization	Research progress
2001	ITU-R	ITU-R model of the atmospheric structure constant for slanted propagation
2005	Patterson	Detection probability of OAM state after scattering
2008	Anguita	Multiplexing crosstalk in the Kolmogorov spectrum
2008	Gbur	Topological variation and conservation in turbulent flows
2009	Tyler	OAM-encoded information loss rate
2009	Wang	Vortex beams spreading in turbulence
2011	Pu	Scintillation index of elliptical vortex beams
2012	Ding	Intensity distribution and broadening of off-axis LG beams
2012	Lukin	Beam spreading of low-order vortex beams in turbulence
2013	Aksenov	Spot drift of LG beams in turbulence
2014	Banakh	LG beam broadening effect
2016	Ge	Beam spreading of vortex beams under different turbulence parameters
2016	Filimonov	Effect of beam and receiver axis-alignment errors on OAM fluctuations
2017	Porfirev	Propagation characteristics of vortex beams in a random aerosol medium
2017	Tian	Probability density of the average OAM mode
2019	Luo	Beam spreading of the vortex beam array and its evolution
2020	Yan	Beam quality of Gaussian and super-Gaussian vortex beams
2022	Lazrek	Propagation characteristics of vChGB in ocean turbulence
2009	Cai	Partially coherent standard and elegant LG beam spreading in free space
2011	Cui	Angular width and transmission factors of partially coherent LG beams in turbulence
2015	Tang	Influence of different parameters and random media on MGSMV beams
2017	Yao	Beam spreading and angle broadening of MGSMV beams induced by the non-Kolmogorov spectrum
2018	Gbur	tvGSM propagation characteristics and OAM density
2021	Hyde	Synthesis and transmission of STOV in free space
2022	Hricha	PCvChGB through the focusing characteristics of a lens system
2009	Haus	Intensity distribution, polarization degree, and scintillation index of VVB
2013	Martinez	Polarization structure evolution of PCRPV
2015	Jiang/Cheng	Polarization characteristics of EPVB after turbulence; influence of linearly polarized butterfly–Gaussian vortex beams
2016	Guo	Characteristic changes of partially PCRPV beams after an ABCD matrix
2016	Eyyuboğlu	Scintillation characteristics of VVB in strong turbulent regions
2018	Tang	Radial polarization MCGSM beams in the non-Kolmogorov spectrum
2020	Joshi	PCVVB polarization variation
2021	Joshi	PCVVB statistical features

(continued)

Table 48.2 (continued)

Year	Person/ organization	Research progress
2022	Manisha/ Arora	Anisotropic PCVVB far-field intensity distribution, V-point polarization vector-beam perturbation
2009	Starikov	Shack–Hartmann wavefront correction
2012	Garanin	SH method for wavefront correction
2012	Leach	SH and GS algorithms used to correct the wavefront of single-mode beams
2007	Jesacher	GS algorithm for wavefront correction
2018	Baranek	Improved GS algorithm for aberration correction
2015	Xie	Wavefront correction of SPGD algorithm based on Zernike polynomial
2018	Xu	Dual wavefront correction based on a liquid-crystal AO technique and PD algorithm
2020	Gao	Turbulence-aberration correction model based on a CNN
2021	Zhan	SH, SPGD, and GS algorithms for underwater-turbulence correction models

spiral spectrum, and Strehl ratio to the optical-transfer and image-resolution functions.

Vortex Beam Separation Measurement

In 2002, Leach et al. proposed a method for OAM state detection based on a Mach–Zehnder interferometer to separate multimode OAM states, but only the odd–even superposition of OAMs could be detected separately [113].

In 2003, Vasnetsov et al. proposed a method to measure the OAM topological charge of a vortex beam using the rotational Doppler effect. They detected the continuous rotation of the interferometric pattern by an experimental method using an off-axis aperture and derived the measured topological charge according to the law of beam rotation [114].

In 2009, Ghai et al. studied the light intensity distribution of a vortex beam after diffraction through a single slit, and the OAM of the vortex beam was determined by observing the direction and magnitude of the curvature of the stripe [115].

In 2014, Silva et al. derived the OAM magnitude and sign of the incident light, based on the distribution rule of diffraction patterns of LG beams, after passing through triangular and square apertures. They found that square apertures could detect topological charges up to 20 orders, which was twice as high as triangular apertures [116].

In 2014, Emile et al. conducted double-slit interference experiments using a vortex beam, and the experimental setup was simple and easy to operate. The direction and magnitude of the movement of the interference fringe could be observed by the naked eye, and the direction and magnitude of the incident vortex beam OAM were obtained [117].

In 2014, Zhou et al. designed a dynamic double-slit interference-detection system by taking two slits on an opaque screen, and alternating the output optical power between high and low by scanning the center of the slit. Using this property, the OAM of the beam could be measured. The scheme was simple to implement, inexpensive, robust to beam incidence deviation and mode impurities, and could measure higher-order OAM modes [118].

In 2016, Fu et al. extended the detection range of Dammann vortex gratings from −12 to +12 to −24 to +24 by synthesizing Dammann vortex gratings with −12 to +12-order spiral phase diagrams. They detected single-mode or superposition-state OAM by the appearance of solid bright spots in far-field diffractive diagrams [119].

In 2016, Li et al. improved the gradually changing-period amplitude grating by proposing a gradually changing-period phase grating. They then introduced a blazing grating and replaced its structural parameters with special trigonometric functions to concentrate the energy at the first-order diffraction level, thus improving the detection accuracy [120].

In 2017, Taira et al. demonstrated for the first time that the off-axis diffraction of an incident LG beam using a simple circular aperture could detect an incident vortex beam. The singularities of the vortex beam were continuously split along the direction perpendicular to the circular aperture motion, and the number and orientation of the split-vortex singularities coincided with the topological charge magnitude and sign of the incident beam, respectively [121].

In 2018, Acevedo et al. numerically simulated and experimented on the Fraunhofer diffraction of OAM using rectangular and pentagonal apertures. It was found that the rectangular apertures could only detect the topological charge magnitude of the vortex beam, and the detection range was only 0–5, while the non-equilateral pentagonal apertures could detect both the topological charge magnitude and sign, and the detection range was 0–20, which was twice as large as that of the triangular apertures [122].

In 2020, Li et al. designed a gradually changing-period spiral-spoke grating (GCPSSG), which combined a spiral phase diagram with an axicon hologram and a gradually changing-period phase grating to detect up to 160 topological charges. The grating was well tolerant to environmental vibrations and beam alignment [123].

In 2017, Kotlyar et al. detected vortex beams using astigmatic transformations, introduced by methods such as cylindrical lenses or tilted spherical lenses, and found that cylindrical lenses were the most effective, with a maximum measured topological charge of 100 [124].

In 2018, Li et al. coupled an incident OAM beam into a single-surface plasmon polarized beam with different splitting angles using an on-chip plasmon nanograting. The magnitude and sign of the OAM could be detected using this method [125].

In 2018, Zhang et al. proposed a convolutional neural network (CNN)-based multilayer representation-learning method to measure multiple OAM states after turbulence, with a detection accuracy of approximately 95.2% in six typical turbulent flows [126].

In 2020, Giordani et al. used enhanced CNN and support vector machine (SVM) algorithms, such as principal component analysis (PCA) and dimensionality reduction (DR) algorithms, which could effectively extract the properties of high-dimensional photon-vector vortex beam systems with a final recognition classification accuracy of more than 90% [127].

In 2021, Rasouli et al. designed a parabolic linear grating to detect the radial index and topological charge of OAM beams. The first-order diffraction level pattern in the simulation and experiments consisted of an $(l + p + 1) \times (p + 1)$ two-dimensional array of optical intensity points. The results showed that vortex beams with a radial index of up to five could be measured [128].

In 2022, Wang et al. proposed the detection of high-order OAM beams and multi-OAM beams using a far-field diffraction fringe generated by a forked grating, while the accuracy and reliability of fringe detection could be improved by increasing the number of bifurcations and offsets [129].

In 2010, Berkhout et al. proposed a coordinate-conversion method that used two static optical elements to effectively separate and measure OAM light states. This method converted a vortex beam containing a spiral phase into a beam with a transverse-phase gradient and then separated the OAM states of single-mode or superimposed vortex beams, according to the distribution of horizontal positions. It finally enabled the separation and measurement of 11 OAM states [130].

In 2013, Mirhosseini et al. improved the coordinate-transformation method proposed by Berkhout by combining the replication-technique (fan-out) algorithm of the beam, so that the lateral phase-gradient distribution after the separation of different modes of vortex light became finer and the overlap between adjacent modes became lower. The separation efficiency increased from 77% to 92%, compared to the normal polar-coordinate transformation, which was beneficial for the efficient separation and detection of OAM superposition states [131].

In 2017, Ruffato et al. designed high-resolution pure-phase diffraction-optics devices and used them for separating and detecting OAM modes with a separable topological charge range of -14 to $+14$, eventually reaching a crosstalk value of -15 dB. The device is expected to be integrated into optical platforms or optical fibers for optical processing [132]. In the same year, Ruffato et al. proposed a compact OAM-mode sorter with logarithmic polar-coordinate transformations, which integrated the coordinate-transformation and phase-correction functions on a phase mask, improving the compactness of the OAM-mode demultiplexer and simplifying the alignment operation [133].

In 2017, Zhao et al. [134] proposed a coordinate-transformation method for separating OAM modes with a radially varying phase, where the OAM topological charge number and radially varying phase determined the horizontal and vertical positions. Simulation and experimental results showed that two to three OAM superposition modes could be separated, and the separation efficiency was improved compared with that in Refs. [130, 131].

In 2018, Chen et al. spirally transformed an OAM beam by representing the phase diagram corresponding to the OAM beam as a spiral line and then converting the collected spiral lines to parallel lines to achieve n-fold spot broadening. Simulations

and experiments showed that this method improved the optical refinement by nearly a factor of three compared to the log-polar transformation and finally separated high-resolution OAM modes [135].

In 2019, Ruffato et al. integrated a beam expander and phase corrector on the same side of a transparent quartz plate to implement a compact OAM-mode sorter. This method enabled OAM-mode separation and detection from −10 to +10 and extended the feasibility to non-paraxial conditions, significantly improving the integration and alignment of the system [136].

In 2019, Fontaine et al. generated and separated 210 LG modes using a multi-plane light converter (MPLC) system consisting of seven phase plates, SLMs, and a reflector [137]. Then, an MPLC inverse-design method was proposed by combining the particle-swarm algorithm and wavefront-matching algorithm to efficiently search the design parameters [138].

In 2021, Huang et al. designed a diffraction-depth neural network (D^2NN) with hybrid OAM-mode generation, detection, and switching functions to achieve full optical signal-processing communication for OAM keying, multiplexing, demultiplexing, and mode switching with a diffraction efficiency and mode-switching purity of more than 96% [139] (Table 48.3).

OAM separation and detection are generally divided into interference [113, 117, 118] and diffraction methods [115, 116, 119–124]. Interference methods include double-slit [117] and dynamic double-slit interference [118], in addition to common plane-wave and spherical-wave interference. Previous interference methods mainly detected OAM beams, based on the orientation and number of interference fringes, whereas dynamic double-slit interference pioneered the use of double-slit diffraction from a rotating opaque screen to detect higher-order OAM beams, based on the power fluctuation of the received interference spot.

Diffraction methods are available for detection using various apertures [115–117, 121, 122] and synthetic gratings [119, 120, 123, 128, 129]. Synthetic gratings are mainly composed of forked [129], gradually changing-period [140, 141], annular [142], Dammann [119], and conical curves [128]. They can be used to measure the performance, which can improve the maximum measured topological charge and diffracted spot-energy distribution. Ultimately, the magnitude and sign of the topological charge number were determined by judging the fringes and orientation of the Hermite–Gaussian-like diffracted spot.

Another class of diffraction methods is called coordinate-transformation methods [130–136]. The coordinate-transformation method separates and detects the OAM beam using two static optical coordinate-transformation and phase-correction elements. It converts the doughnut-type spot into a striped rectangular spot and obtains the mode distribution of the incoming OAM beam, based on the calculation of the spot position and coordinate scale. On this basis, the accuracy of the diffracted spot can be improved using fan-out algorithms [131]. Alternatively, two phase diagrams can be synthesized into a single-phase mask to improve the detection alignment and system integration.

Finally, there is currently a multilayer light-diffraction neural network that integrates generation detection and transformations [137–139]. The number of

Table 48.3 Domestic and foreign research progress in vortex beam separation and detection

Year	Researcher	Research progress
2002	Leach	Separation and detection of odd–even state superposition by Mach–Zehnder interferometer
2003	Vasnetsov	Rotating Doppler effect
2009	Ghai	Single slit
2014	Silva	Triangular and square apertures
2014	Emile	Double-slit interference
2014	Zhou	Dynamic double-slit interference
2016	Fu	Synthetic Dammann grating
2016	Li	Improved gradually changing-period phase grating
2017	Taira	Simple circular aperture
2018	Acevedo	Square and pentagonal apertures
2020	Li	GCPSSG
2017	Kotlyar	Cylindrical lens
2018	Li	Plasma metasurface
2018	Zhang	CNN multilayer representation-learning method
2020	Giordani	CNN combined with SVM enhancement algorithm
2021	Rasouli	Parabolic linear grating
2022	Wang	Fork-shaped grating
2010	Berkhout	Coordinate-transformation method
2013	Mirhosseini	Coordinate-transformation method of fan-out algorithm
2017	Ruffato	Compact, high-resolution OAM mode sorter
2017	Zhao	Coordinate-transformation method of radial variant phases
2018	Chen	Spiral-transformation method
2019	Ruffato	Compact OAM-mode classification system for log-polar transformation
2019	Fontaine	Particle-swarm algorithm combined with MPLC
2021	Huang	Five-layer diffraction-depth neural network

measurable separated topological charges has increased linearly, and the entire detection system is moving toward high performance, more compactness, and smaller integration.

OAM Technology in Optical Wireless Communication Systems at Xi'an University of Technology

Ke et al. at Xi'an University of Technology have been systematically researching optical wireless communication technology since 2002. They have achieved good results through exploring and researching technologies in signal modulation, error-correction codes, demodulation, wavefront correction, automatic acquisition, tracking, and pointing, optical-antenna design, optical-fiber coupling, atmospheric turbulence transmissions, and system performance [6]. As exploiting existing dimensional resources has reached a bottleneck, the optical communication capacity crisis has

become an urgent problem. The OAM-structured optical field with spatially varying phases has attracted increasing attention and input from scholars.

Since 2008, Ke et al. have studied OAM technology in optical wireless communication. They have conducted simulation and experimental research on OAM beam generation, transmission characteristics, crosstalk analysis, wavefront correction, separation and detection, and application, according to various aspects of OAM-multiplexed communication systems, and made good progress [143]. The research-content structure diagram is shown in Fig. 48.3.

In 2015, Li et al. generated fractional-order vortex beams and their arrays with a topological charge spacing of 0.5, using fractional-order forked gratings and their superposition. They found that the fractional-order vortex beams were stronger than the integer-order optical intensity distribution, which was beneficial for long-distance transmissions [144].

In 2018, Ge et al. experimentally investigated using few-mode fibers to generate an ordered LG beam. The highest excitation efficiency of 18.39% was produced for a low-order vortex beam, when the fiber incidence angle was adjusted to 2.895° [145].

In 2019, Chen et al. improved the inverted-parabolic fiber structure based on Ref. [142] by adding a low refractive index layer to increase the effective refractive index difference. It generated vortex beams of up to the third order [146].

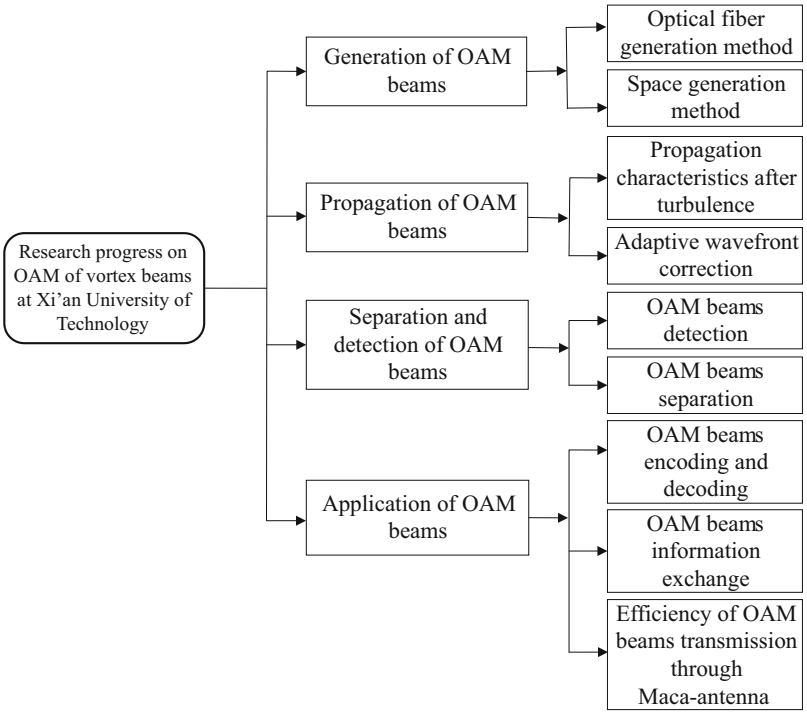


Fig. 48.3 OAM technology research-content structure diagram

In 2020, Wang et al. improved the photonic-crystal fiber (PCF) structure by replacing parts of the circular air holes with rectangular air holes to ensure a high refractive index difference between the cladding and the fiber core, and prevent the leakage of higher-order modes to the cladding. This PCF structure generated and supported the transmission of 50 OAM modes; that is, it had a topological charge $l = 1$–13 [147].

In 2017, Xu et al. used the superposition of forked gratings to generate equal and unequal opposite-sign petal-like superimposed vortex beams, with the number of split petals being the sum of the absolute values of the topological charges of the two beams [148]. In the same year, Xue et al. proposed the superposition of spiral phase diagrams to generate a superimposed vortex beam with the same number of circular aperture splittings. The number of bright spots was equal to the absolute value of the difference between the topological charge numbers of the two beams [149].

In 2018, Shi et al. studied the superposition state generation of high-order radial LG beams, based on Ref. [32], and analyzed the effect of different parameters on the superposition state [150].

In 2019, Zhao et al. simulated the superposition state generation of multiple LG beams of the same or different signs in an arithmetic sequence in the on-axis and off-axis cases and studied its distribution law [151].

In 2020, Zhang studied the generation of orthogonal multi-ring superposition-state vortex beams through simulations and experiments [152].

In 2008, Lv et al. studied the scattering of a single spherical particle with an OAM beam by analogizing the LG beam to a Gaussian-beam model. They obtained the distribution pattern of the scattering intensity and scattering cross section, based on the generalized Mie theory and a local approximation [153].

In 2009, Lv et al. proposed a high-order Bessel diffraction-free beam coding and decoding method for optical OAM by converting LG beams into high-order Bessel diffraction-free beams [154].

In 2009, Yang et al. studied the probability distribution and variation of the single-photon average OAM of high-order elliptical Hermite–Gaussian beams and found that the average OAM varied with the magnitude of different material–matter interactions [155].

In 2011, Lv et al. used the complex Gaussian function expansion method to analyze the effect of the obscuration ratio on the transmission characteristics of LG beams of different wavelengths, using a physical model of LG beams passing through two different aperture optical systems. They found that the effect of the obscuration ratio on the LG beams was greater as the lens-aperture spacing increased and the wavelength became longer [156].

In 2011, Lv et al. used the von-Kármán turbulence spectrum to establish a two-photon entanglement degree model. They analyzed the effects of the transmission distance, turbulence intensity, topological charge number, wavelength, and beam waist on the entangled photons and the entanglement degree [157].

In 2013, Wang et al. investigated the joint detection probability of OAM-entangled photon pairs generated by the conversion process under

spontaneous parameters and found that the beam waist radius and angular and radial indices affected the joint detection probability of the entangled photon pair [158].

In 2014, Yang et al. investigated the transmission characteristics of LG beams on an atmospheric-turbulent slanted path and obtained the spiral spectrum distribution at different turbulence intensities, beam waist radii, topological charge numbers, zenith angles, and receiving apertures [159].

In 2015, Guo et al. designed an eight-stage structure phase-information coding and decoding method to improve the capacity and efficiency of OAM communication [160].

In 2018, Ning et al. studied the optical intensity and phase distribution of multiplexed OAM beams at different turbulence intensities and transmission distances, while analyzing their spiral spectra to obtain the system BER at different turbulence intensities [161].

In 2018, Wang et al. studied the diffraction-transmission characteristics of an LG beam after passing through a Maksutov–Cassegrain antenna (Maca-antenna) optical system. They found that the larger the topological charge number, the higher the emission efficiency of the LG beam after passing through the antenna; it increased to more than 80% [162].

In 2020, Wang et al. analyzed the transmission characteristics of Laguerre–Gaussian–Schell mode vortex beams (LGSMV) in non-Kolmogorov turbulence and derived the far-field light intensity expressions, while obtaining the phase singularity and light intensity distribution variations for different turbulence and beam parameters [163].

In 2020, Chen et al. investigated the information exchange, mode switching, and channel reconfiguration under different OAM modes using simulations and experiments [164].

In 2021, Li et al. analyzed the crosstalk suppression of OAM-MIMO systems in turbulence and compared the improvement of the constellation diagram and BER diagram of the system under the multiuser kurtosis (MUK) modified constant modulus algorithm (MCMA) and cross-correlation (CC)-MCMA algorithm [165].

In 2017, Wang et al. studied the normalized light intensity distribution of partially coherent LG beams under Kolmogorov turbulence and obtained the variation curves of the light intensity distribution with different parameters [166].

In 2019, Wang et al. analyzed the transmission characteristics of partially coherent vector vortex beams—that is, electromagnetic Gaussian–Schell mode vortex beams (EGSMV)—in non-Kolmogorov turbulence and obtained the variation of the polarization direction angle and degree under different parameters [167].

In 2018, Wang et al. investigated adaptive wavefront sensorless correction for distorted vortex beams using the SPGD algorithm. They found that this algorithm could correct wavefront distortions for both single-mode and multimode OAM beams; however, the SPGD algorithm corrected single-mode OAM beams better than it did multimode [168].

In 2020, Cui et al. used a wavefront-free distortion correction technique, based on the PD algorithm for vortex beams, and verified the feasibility by simulation and

experiment. They found that the correction effect of vortex beams improved with the decrease of topological charge number [169].

In 2018, Wang et al. compared and analyzed the far-field diffraction characteristics of an LG beam after passing through circular, square, triangular, single-slit, and double-slit apertures. They found that the topological charge number magnitude and sign of the incident beam could be judged based on the number of spots or stripes in the diffracted light field and their orientation [170].

In 2018, Xie et al. used a variety of gradually changing-period gratings and annular gratings to detect LG beams and proposed that the detection performance of these gratings could be improved by using phase correction diagrams and fan-out phase diagrams. It eventually improved the maximum value of the detectable topological charge number [8].

Generation of Vortex Beams Carrying OAM

To achieve better OAM-multiplexed communication, it is necessary to first generate a vortex beam carrying OAM. Common methods for generating vortex beams are fiber generation and space generation.

Optical-Fiber Generation Method

Vortex beams generated by spatially structured devices cause a certain amount of energy loss. Using optical fibers to generate vortex beams can reduce the light energy loss when passing through structured devices; thus, the conversion efficiency of the fiber-generation method is higher.

Few-Mode Fiber-Generation Method

A few-mode fiber (FMF) is an optical fiber with only fundamental and low-order modes. The HI1060 few-mode fiber was selected for the experiment, and the relationship curve between the incident angle and excitation efficiency is shown in Fig. 48.4a. In the experiment, when the incident angle was controlled between 2.895° and 5.729°, the fundamental mode of the first-order vortex light could be formed in the few-mode fiber. The excitation efficiency of the first-order vortex light was maximized when the incident angle was 2.895°, and its excitation efficiency was 18.39%. Figure 48.4b shows a diagram of the experimental setup for vortex light generation in the few-mode fiber.

Figure 48.5a, b show the simulated and experimental diagrams of the first-order vortex-mode spot in the fiber, respectively. Figure 48.5c, d show the simulated and experimental spot diagrams of the LP_{11} mode, respectively. The LP_{11} mode is formed because of the degeneracy between the multiple vector modes in the fiber.

Improved Inverted-Parabolic Fiber Structure

To increase the effective refractive index difference between the core and cladding, a low refractive index layer is added to improve the few-mode fiber structure into an inverted-parabolic distribution type so that it can contain a larger number of modes.

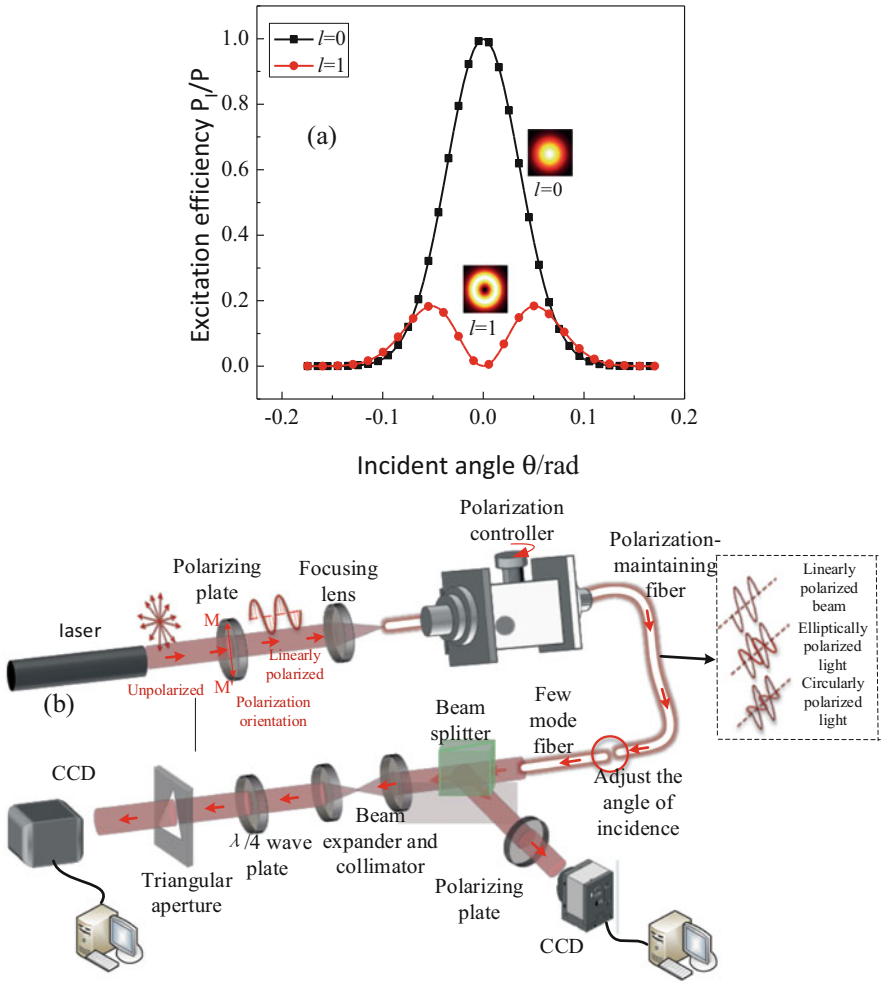


Fig. 48.4 Few-mode fiber experiment. (**a**) Relationship between the incident angle and excitation efficiency of a few-mode fiber; (**b**) experimental-setup diagram of a low-order vortex beam generated by a few-mode fiber [171]

In Fig. 48.6a, the refractive index distribution curves of the improved and original fiber structures are shown in black and red, respectively.

The relationship between the effective refractive index and wavelength is shown in Fig. 48.6b. The diagram shows that the fiber structure can contain nine modes in the fiber core refractive index range and transmit up to third-order OAM modes. The light intensity and phase distribution of the generated third-order OAM modes are shown in Fig. 48.7.

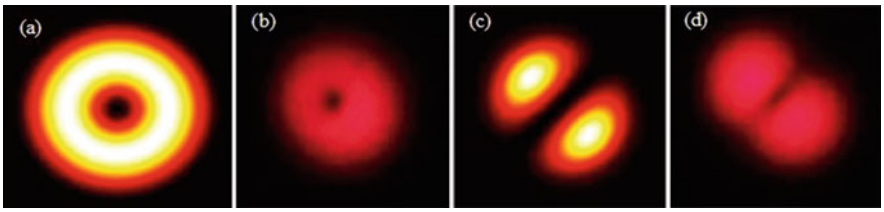


Fig. 48.5 Intensity of few-mode fiber in first-order mode. (**a**) Simulation of first-order mode; (**b**) experiment of first-order mode; (**c**) simulation of LP_{11}; and (**d**) experiment of LP_{11} [171]

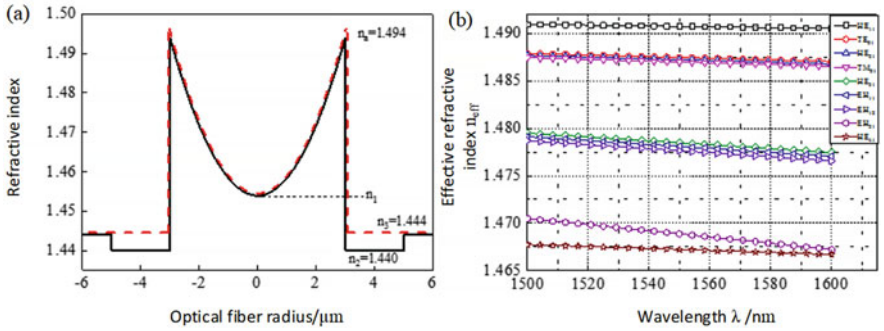


Fig. 48.6 Refractive index curve of the fiber structure. (**a**) Refractive index distribution of improved fiber structure; (**b**) relationship between vector effective refractive index and wavelength [172]

Improved Photonic-Crystal Fiber Structure

By combining rectangular and circular air holes, the cladding layer has a larger air-filling ratio, which increases the refractive index difference with the annular high refractive index layer and prevents the leakage of higher-order modes into the cladding layer. Moreover, this structure is less deformable and more stable. The cross section and quarter cross section of the modified PCF structure are shown in Fig. 48.8a.

The odd and even modes of the HE modes are superimposed to generate the optical field and phase diagram of the OAM modes, as shown in Fig. 48.8b. The first and second rows are the HE_{21}^{even} and HE_{21}^{odd} modes, respectively. The HE_{41}^{even} and HE_{41}^{odd} modes are superimposed to generate the optical field and phase-distribution diagram of the OAM_{11}^{+} modes and OAM_{31}^{+} modes. Based on the OAM mode-generation process and analyzing the vector modes in the PCF, it was concluded that this PCF could support the transmission of $l = 1$ to 13 vector-mode topological charges in the 850-nm band range. Thirteen $OAM_{\pm l,m}^{\pm}$ mode groups were synthesized, with a total of 50 OAM modes. Compared to the 26 modes in Ref. [174] and 58 modes in Ref. [175], the band is wider and the number of modes is larger (Table 48.4).

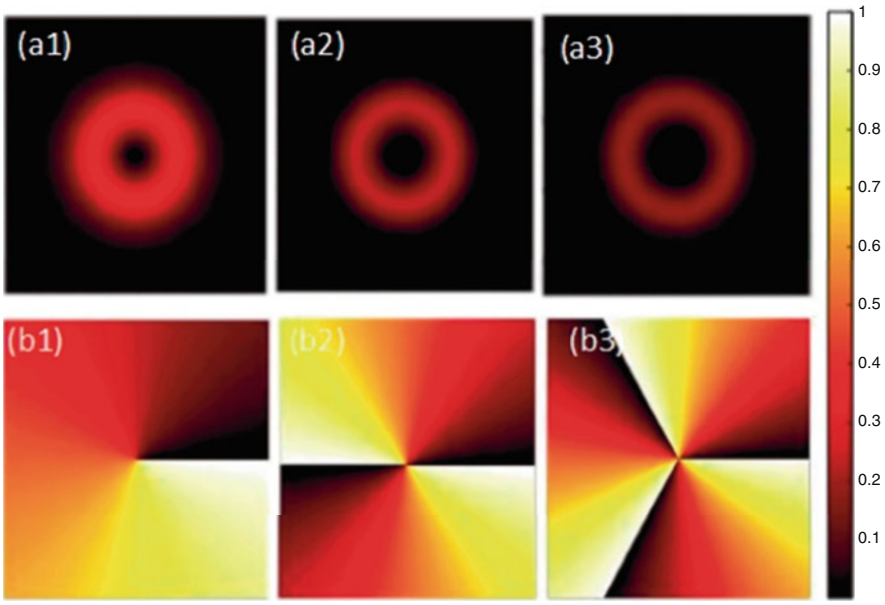


Fig. 48.7 Intensity and phase distribution of vortex beams generated by improved fiber. (a1) OAM_{11}; (b1) OAM_{21}; (c1) OAM_{31}; (a2) distribution of OAM_{11}; (b2) distribution of OAM_{21}; (c2) distribution of OAM_{31} [172]

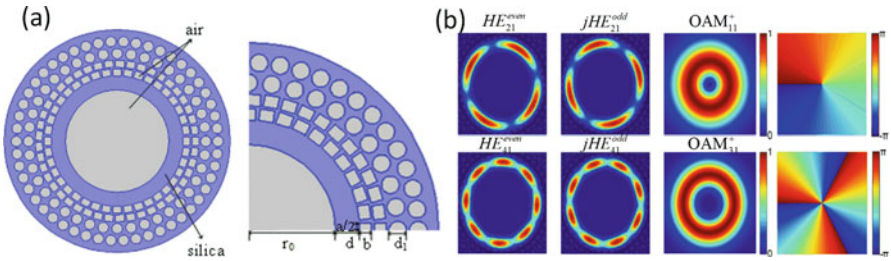


Fig. 48.8 Improved OAM in PCF. (**a**) Improved structural PCF cross section; (**b**) OAM-mode generation in this PCF [173]

Spatial-Generation Method

Generally, direct-generation methods, mode-conversion methods, spiral phase plates (SPPs), and CGH are used to generate vortex beams using a spatial structure. Here, we mainly utilize the CGH method and use a programmed holographic grating or phase diagram to load the spatial light modulator for modulation by a computer. Then, an OAM vortex beam of the desired structure can be obtained.

Table 48.4 Summary of optical fiber vortex beam generation

Fiber-generation method	Structure principle	Performance analysis
FMF	Adjust incident angle of FMF	Excitation efficiency of 18.39% with±1order vortex beam
Inverted-parabolic fiber structure	Addition of low refractive index layer	Produces first- to third-order vortex beams
Improved photonic crystal fiber	Rectangular air holes with circular air holes	Produces first- to third-order vortex beams in the 850-nm range

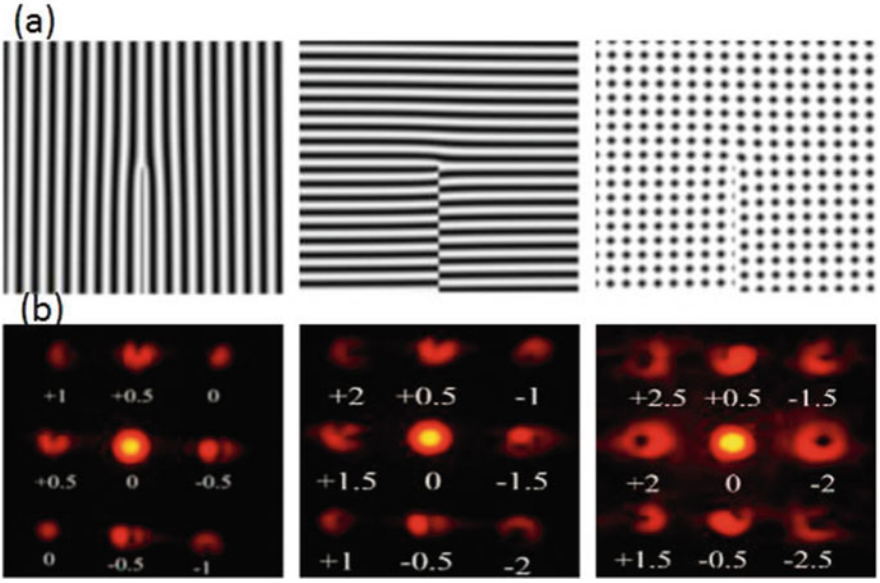


Fig. 48.9 Fractional vortex beam generation. (**a**) Fractional composite fork-grating generation; (**b**) fractional vortex beam array [176]

Fractional-Order LG Beam Generation

The vertical dislocation $l = 0.5$ and a transverse fractional-order forked grating are superimposed to obtain a fractional-order composite forked grating, as shown in Fig. 48.9a. Figure 48.9b shows the experimental diagram of the fractional-order vortex beam array; the notch distribution becomes larger as the value of the topological charge fraction increases.

LG Beam Superposition State Generation

First, different superposition states of vortex light can be generated by superposition gratings, as shown in Fig. 48.10a, which are the superposition of $l = 2$, $l = -3$, $l = 3$, and $l = -3$ fork gratings, respectively. The Gaussian beam strikes on top of the superimposed grating to produce a superimposed OAM vortex beam, as shown in Fig. 48.10b in simulation and experimentally.

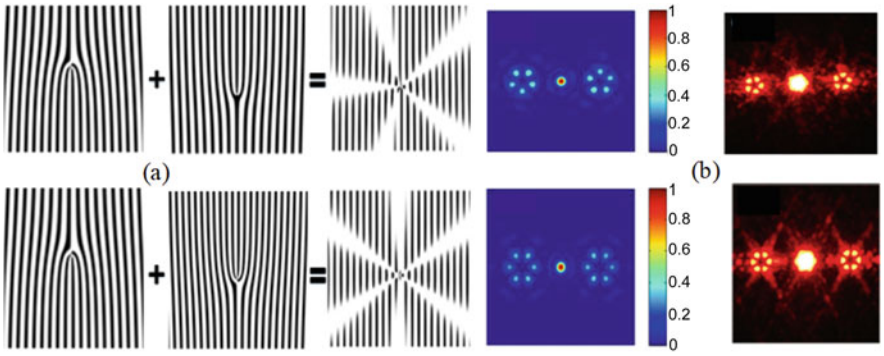


Fig. 48.10 Superposition state preparation of an LG beam. (**a**) Superposition of fork gratings; (**b**) LG beam superposition state with different topological charges [177]

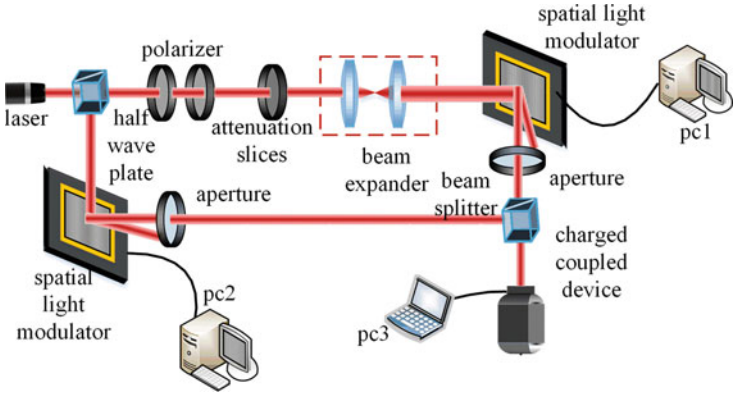


Fig. 48.11 Diagram of experimental setup for superposition state generation

The experimental setup for interferometric superposition state generation is shown in Fig. 48.11. It mainly uses two SLMs loaded with two forked gratings for coaxial superposition interference, but it is also possible to load the superposition grating onto one SLM to achieve efficient coherent superposition.

In addition to the LG beam superposition state prepared using the superposition of forked gratings, spiral phase diagrams can also be used. The superimposed phase diagrams of $l = 1, 5$ and $l = 2, 4$ are shown in Fig. 48.12 (a1) and (b1). The experimental setup was similar to that shown in Fig. 48.11.

The generated superimposed vortex beams are shown in Fig. 48.12(a2–a3) and (b2–b3), which are the simulated and experimental diagrams corresponding to the

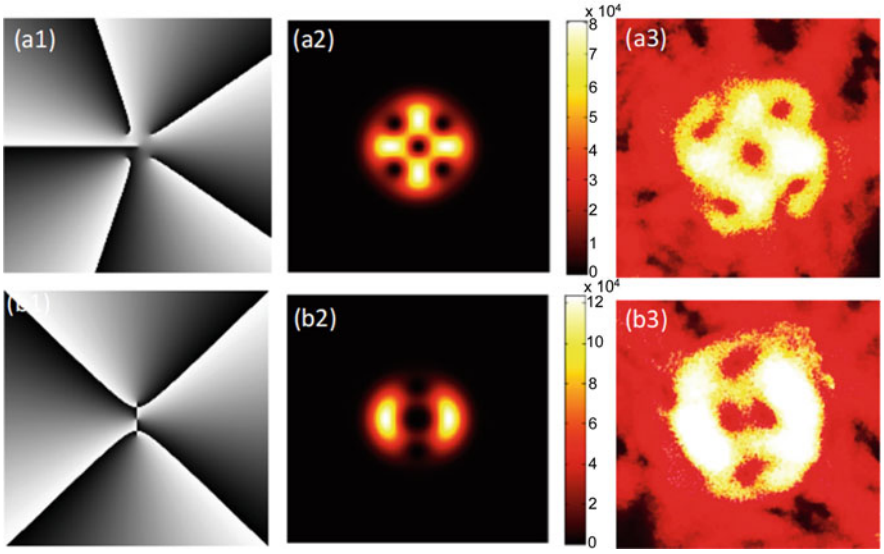


Fig. 48.12 Superimposed LG beams generated by phase patterns. (a1–a3) $l = 1, 5$ phase-diagram superposition and corresponding simulation and experimental diagrams; (b1–b3) $l = 2, 4$ phase-diagram superposition and corresponding simulation and experimental diagrams [178]

superimposed state of the same number, respectively. At this point, a circular splitting spot is formed and the number of bright spots is $N = |l_1 - l_2|$.

Another type of orthogonal superposition state generation takes advantage of the orthogonality between different OAM modes to orthogonally superimpose two vortex beams with different topological charge numbers; using this double-ring structure for information transmission can reduce inter-mode crosstalk.

The simulated (a1–a3) and experimental (b1–b3) diagrams of the double-ring vortex beam, after the orthogonal superposition of the same number of l values, are shown in Fig. 48.13. It has the same experimental setup as shown in Fig. 48.11. The only difference is that the outgoing beam needs to pass through a polarizer after the SLM reflection to control its angle, so that the two vortex beams finally incident on the polarization combiner are in an orthogonal state.

Finally, a multi-LG beam superposition state was generated, and the light field distribution of the superposed vortex beams with different topological charge numbers was obtained according to the theoretical calculation analysis, as shown in Fig. 48.14. The tolerances shown in Fig. 48.14(a1–d1) and (a2–d2) are the simulation results of the same number of multi-LG mode superpositions for $\Delta = 1$ and $\Delta = 2$, respectively.

As can be seen from the figure, because the coaxial superimposed light intensity diagram shows bright and dark spots, Fig. 48.14(a1–d1) all show one bright spot. There is a singularity in the center of the superimposed light field with dark spots on the outer side, and the number of dark spots is $1 \times (n - 1)$. Figure 48.14(a2–d2) all show two bright spots, and the number of dark spots on the outer side is $2 \times (n - 1)$.

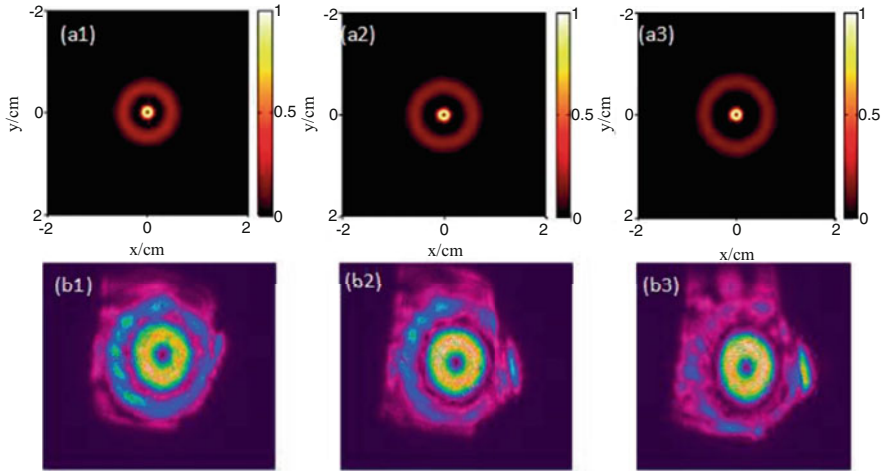


Fig. 48.13 Simulation and experimental diagram of orthogonal superposition states of the same sign l value. (a1–b1) $l = 1, 5$; (a2–b2) $l = 1, 7$; (a3–b3) $l = 1, 9$ [179]

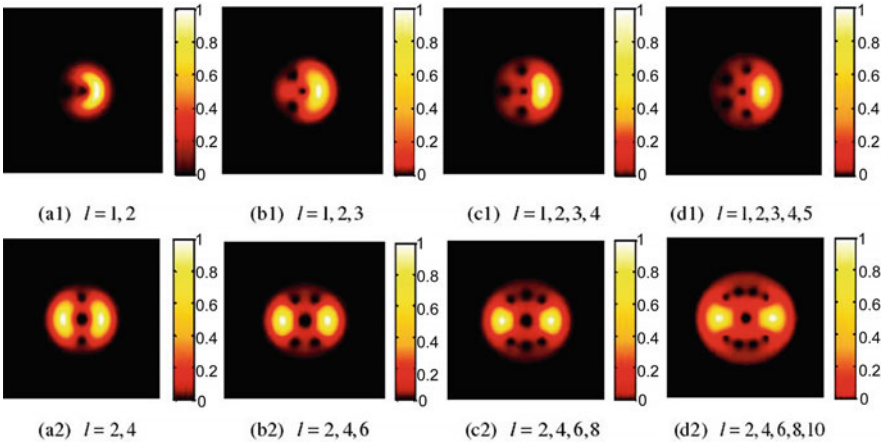


Fig. 48.14 Light intensity distribution diagram after multiple superimposed LG beams of the same sign [180]

The result is the same for different numbers, except that the spots change from circle-like to petal-like at this time.

High-Order Radial LG Beam Superposition State Generation

A higher-order radial LG beam superposition state is used in OAM-multiplexed communication systems to improve the capacity, spectrum utilization, and received power of the communication system. Figure 48.15(a1–a3) show the superimposed light fields of LG_0^3 and LG_1^3, LG_0^3 and LG_2^3, LG_0^3 and LG_3^3, respectively; the light

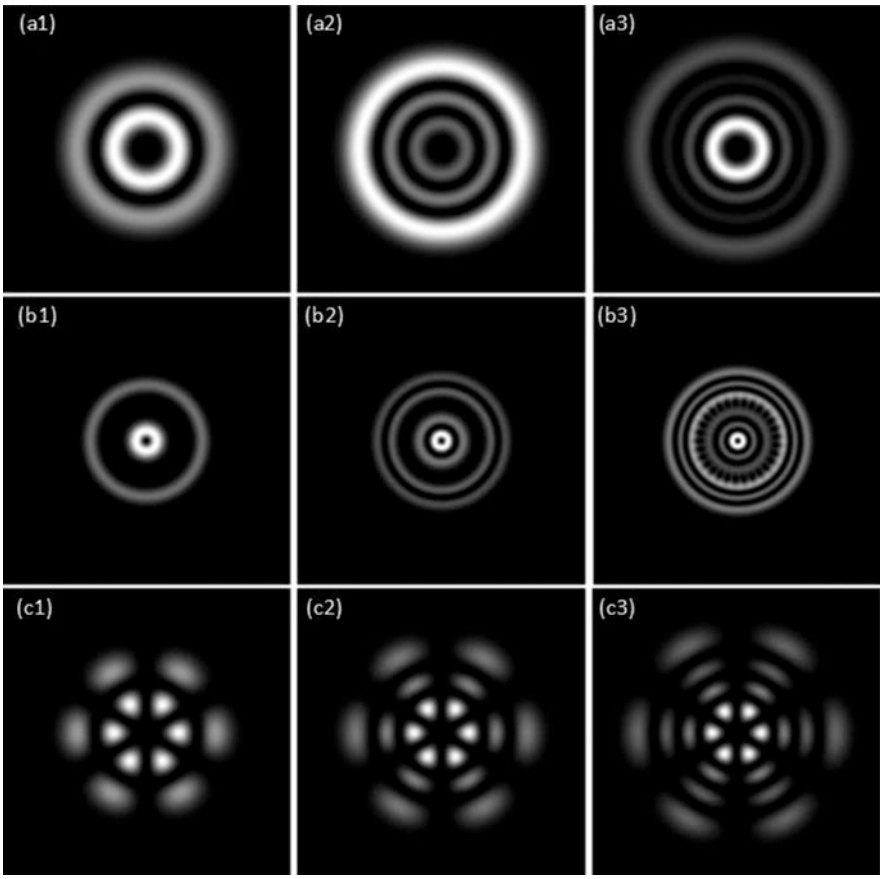


Fig. 48.15 High-order radial LG beam superposition. (a1–a3) Same topological charge number and different radial index; (b1–b3) same radial index and different topological charge number; (c1–c3) same radial index and topological charge with opposite numbers of each other [181]

intensity distribution is in the form of multiple bright rings and their bright ring numbers $\sigma = \max\{p_1, p_2\} + 1$.

Figure 48.15(b1–b3) show the superimposed light fields of LG_0^1 and LG_0^{30}, LG_1^1 and LG_1^{30}, LG_2^1 and LG_2^{30}, respectively. The number of bright rings $\sigma = n(p + 1)$ and n represents the number of coaxially superimposed LG beams. Figure 48.15(c1–c3) show the superimposed light fields of LG_1^3 and LG_1^{-3}, LG_2^3 and LG_2^3, LG_3^3 and LG_3^3, respectively. They are distributed in the shape of a petal with a symmetric distribution of layer circles, and the total number of bright "petals" is $(p + 1)|l_2 - l_1|$ (Table 48.5).

Vortex Beam OAM Transmission

Here, the transmission characteristics of LG beams in oblique-range turbulence and partially coherent Gaussian–Schell mode (GSM) vortex beams in turbulence are

Table 48.5 Summary of spatial vortex beam generation

Spatial-generation method	Shape and properties of light spots
Fractional grating	Fractional vortex beam array
Grating superposition	The number of petal-like spots is superimposed with different signs $N = \|l_1\| + \|l_2\|$
Phase superposition	The same sign is superimposed as a circular split spot $N = \|l_1 - l_2\|$
LG beam orthogonal superposition	Double or triple-ring superposition spot
Arithmetic distribution of multiple LG beams with the same number superposition	Number of dark spots on the outside of circular-split light spots $\Delta \times (n-1)$
Double high-order radial LG beam superposition	Multiple rings $\sigma = \max\{p_1, p_2\} + 1$, $\sigma = n(p+1)$, many-petal shape $(p+1)\|l_2 - l_1\|$

discussed first. They are followed by an analysis of the crosstalk characteristics of multiplexed vortex beams in turbulence, the transmission characteristics of partially coherent vortex beams in Kolmogorov and non-Kolmogorov turbulence spectra, and the evolution and transmission characteristics of vector partially coherent vortex beams and their arrays in turbulence. Finally, the crosstalk-suppression performance of an OAM-MIMO system under turbulence with different algorithms is analyzed.

Transmission Characteristics After Turbulence

After the LG beam is transmitted through slanted atmospheric turbulence, it causes OAM-mode dispersion and inter-mode crosstalk, as shown in Fig. 48.16(a1–d3), for the distribution of the spiral-spectral components with the transmission distance for different near-ground refractive index structure constants, incident light wavelengths, receiving apertures, zenith angles, and topological charge numbers, respectively.

It can be observed that the OAM divergence slows as the transmission distance increases. In contrast, the larger the parameters, except for the opposite wavelength, the stronger the OAM spiral spectrum dispersion. Therefore, the selection of a long wavelength, small receiving aperture, small zenith angle, and small topological charge number can reduce the OAM dispersion.

Subsequently, the transmission characteristics of the GSM vortex beam under Kolmogorov turbulence were investigated. Figure 48.17a–f show the normalized light intensity distribution curves for different turbulence intensities, transmission distances, topological charge numbers, coherence lengths, beam waist radii, and wavelengths, respectively.

It can be seen that the stronger the turbulence intensity, the longer the transmission distance, the smaller the topological charge number, the smaller the coherence length, and the larger the beam waist width, the more easily the beam is converted into a Gaussian distribution, and the more influenced by turbulence it is; whereas, the wavelength has no effect.

The OAM beam is affected by turbulence when it is transmitted in free space, which leads to inter-mode crosstalk and OAM-mode dispersion. This increases the BER and reduces the performance of the communication system. Figure 48.18

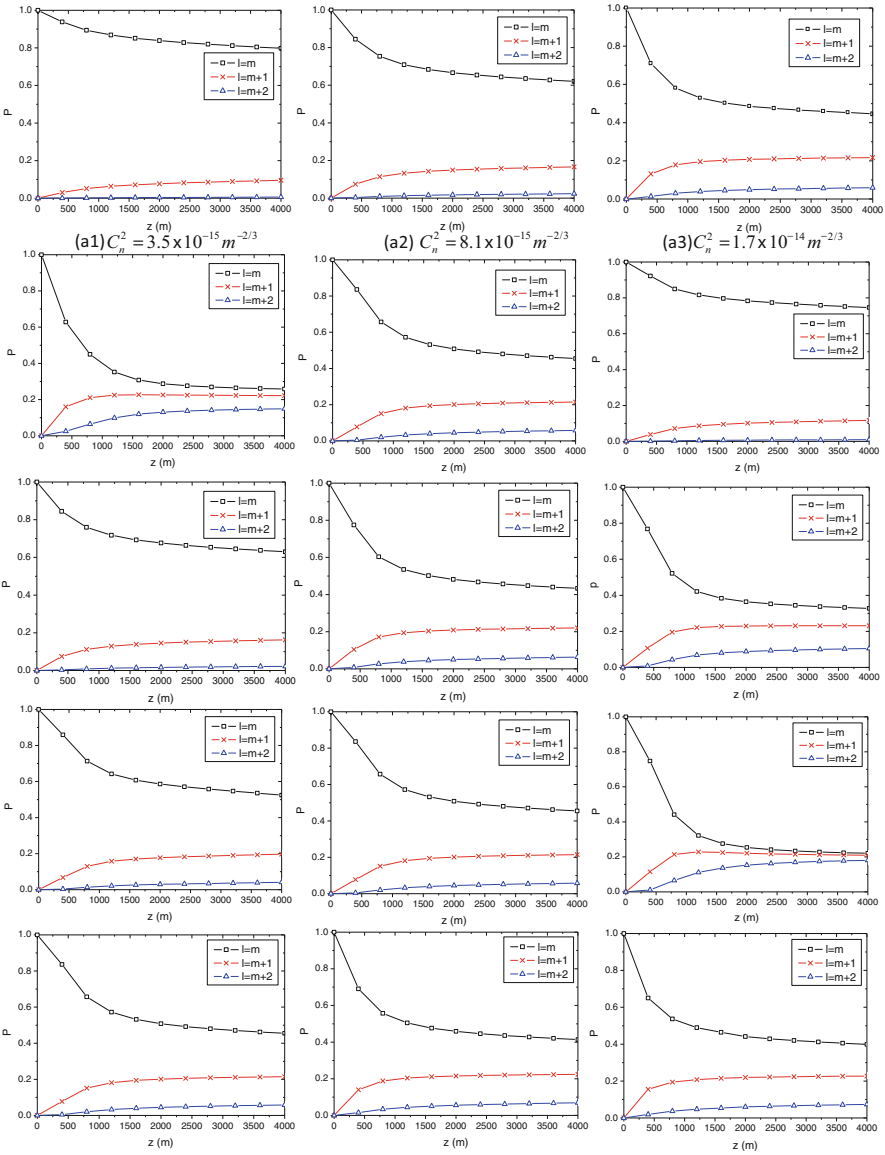


Fig. 48.16 Distribution of helical-spectrum components with propagation distance under different parameters. (a1–a3) Near-ground refractive index structure constant: $C_n^2(0) = 3.5 \times 10^{-15} m^{-2/3}$, $8.1 \times 10^{-15} m^{-2/3}$, $1.7 \times 10^{-14} m^{-2/3}$; (b1–b3) wavelength: $\lambda = 351nm$, $632nm$, $1068nm$; (c1–c3) receiving aperture: $R = 2$cm, 3cm, 4cm; (d1–d3) zenith angle: $\alpha = \pi/4$, $\pi/3$, $\pi/2$; (e1–e3) OAM index: $m = 3$, 10, 20 [159]

(a1–a4) and (b1–b4) show the spiral spectrum distribution of the multiplexed OAM beams of $l = 2$, 4, 6, 8 at different turbulence intensities and transmission distances, respectively.

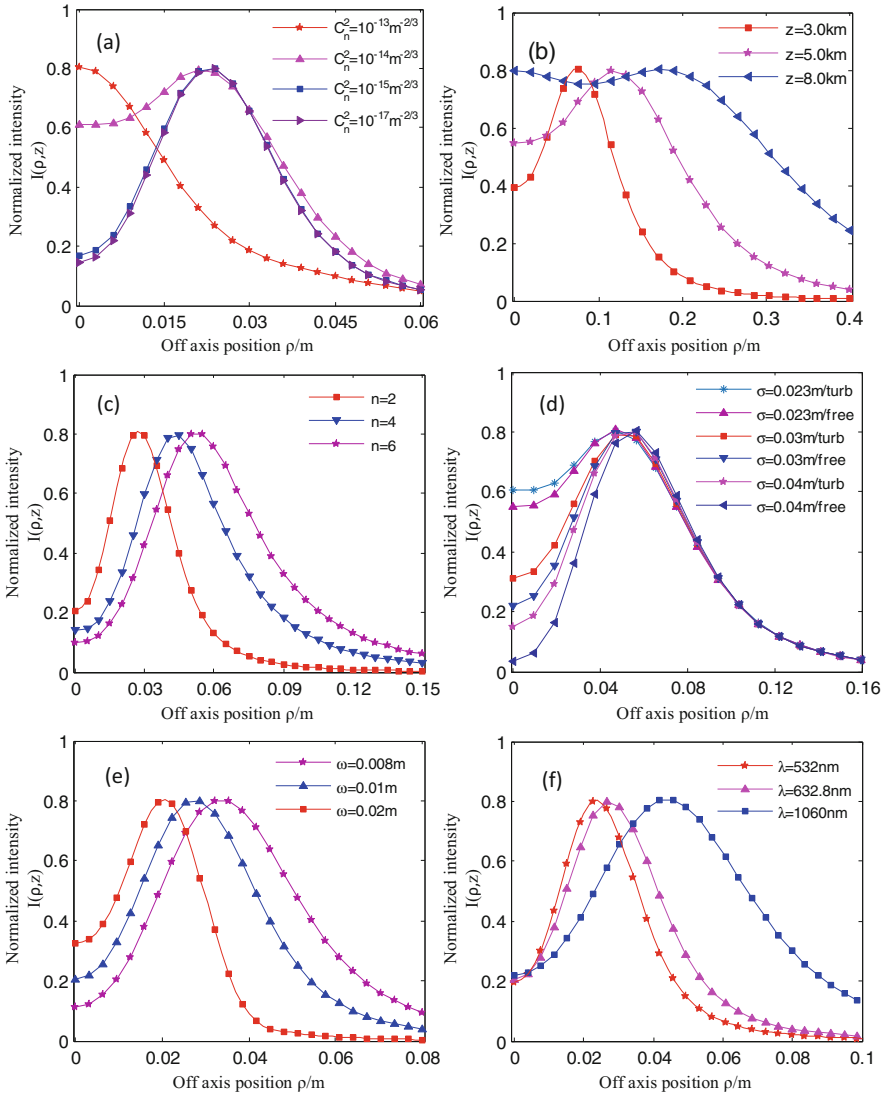


Fig. 48.17 Normalized intensity distribution of GSM vortex beams under different parameters. (**a**) Turbulence intensity; (**b**) transmission distance; (**c**) topological charge number; (**d**) coherence length; (**e**) beam waist radius; (**f**) wavelength [182]

With the increase in turbulence intensity and transmission distance, the mode dispersion degree of each OAM state gradually becomes severe; however, it can still be discerned by the relative power. The dispersion is the most serious at (a4) and (b4), where each OAM mode is dispersive to other modes and the original beam information can no longer be identified.

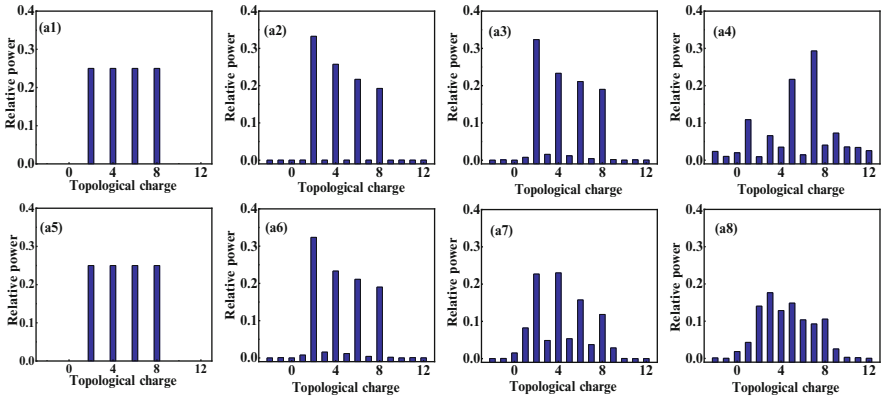


Fig. 48.18 Spiral spectrum distribution of OAM-multiplexed beams under different turbulence intensities and transmission distances. (a1) no turbulence; (a2) $C_n^2 = 1 \times 10^{-17} m^{-2/3}$; (a3) $C_n^2 = 1 \times 10^{-15} m^{-2/3}$; (a4) $C_n^2 = 1 \times 10^{-13} m^{-2/3}$; (a5) $z = 0$ m; (a6) $z = 1000$ m; (a7) $z = 3000$ m; (a8) $z = 5000$ m [183]

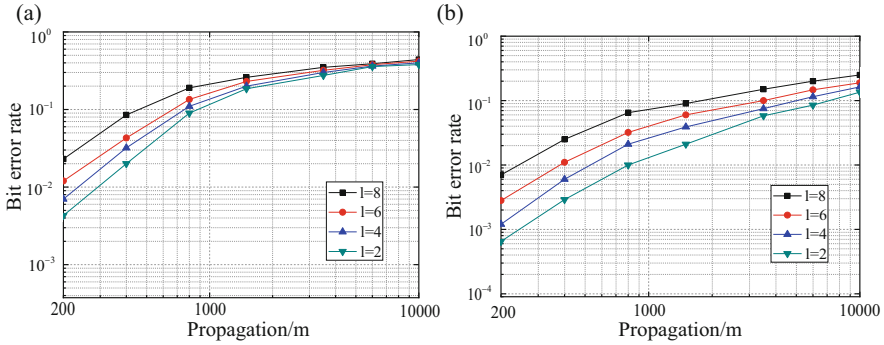


Fig. 48.19 Bit error rate curve of the system under different turbulence intensities (**a**) $C_n^2 = 5 \times 10^{-14} m^{-2/3}$ (**b**) $C_n^2 = 1 \times 10^{-17} m^{-2/3}$ [183]

In Fig. 48.19, as the transmission distance increases, the BER gradually increases, and when the transmission reaches 10,000 m, the BER stabilizes at about 0.4 because the noise reaches saturation. The turbulence intensity decreases to about 0.2 when it becomes smaller, while the larger the topological charge in the multiplexed beam, the higher the BER.

Beam arrays not only increase the emitted power of the light source but also suppress turbulence. This is a hot research topic in high-energy laser synthesis and space-optical communication. Figure 48.20a–c show the light intensity distribution of a radial partially coherent vortex beam array with radial radii of $r_n = 5w_0$, $r_n = 7w_0$, and $r_n = 10w_0$, respectively, after non-Kolmogorov turbulent transmissions of 0.5, 1.5, 2, 2.5, and 3 km.

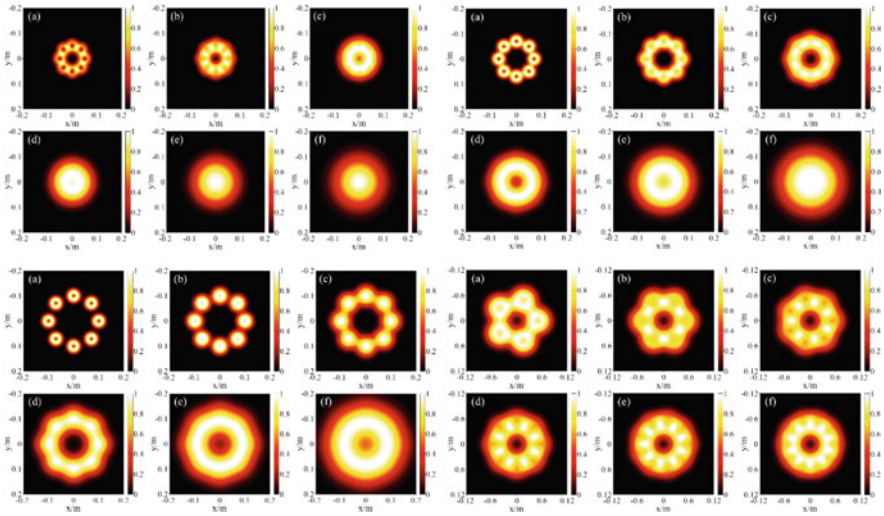


Fig. 48.20 Light intensity distributions of radial partially coherent vortex beam arrays transmitted through non-Kolmogorov turbulence for 0.5 km, 1 km, 1.5 km, 2 km, 2.5 km, and 3 km under different parameters. (**a**) $r_n = 5w_0$; (**b**) $r_n = 7w_0$; (**c**) $r_n = 10w_0$; (**d**) $N = 5$–10 [184]

With the increase in transmission distance, the beam of each array gradually transforms into a Gaussian-like beam distribution. The larger the radial radius, the longer the transmission distance required to convert to a Gaussian distribution. In the same turbulent case, a larger radius is required for conversion to a Gaussian distribution. Figure 48.20d shows the light intensity distribution after transmission with different numbers of arrays at $r_n = 5w_0$ and $z = 1$ km. This illustrates that when the number of radial arrays is greater, the light intensity distribution is closer to the circular distribution, that is, a shorter distance is required to convert to a Gaussian distribution.

Figure 48.21a–d show the light intensity distribution and curves of a partially coherent vortex beam array after non-Kolmogorov turbulent transmissions at different wavelengths, beam waist radii, outer scales, and turbulence intensities, respectively. From these figures, it can be seen that the transmission distance required to convert to a Gaussian distribution is shorter for larger wavelengths, beam waists, and turbulence intensities, which are more influenced by turbulence. The change in the outer scale had almost no effect on the light intensity distribution. The array beam was most affected by turbulence when the generalized exponential factor was $\alpha = 3.15$, as shown in Fig. 48.21d.

The partially coherent vortex beam in a Laguerre–Gaussian–Schell-mode vortex (LGSMV) beam is a new type of vortex field consisting of a Laguerre–Gaussian–Schell (LGSM) beam after spiral phase modulation. Figure 48.22a, b show the light intensity distribution curves of the LGSMV beam with $l = 3$ after non-Kolmogorov turbulent transmissions at the generalized exponential factors $\alpha = 3.1$ and $\alpha = 3.5$ for different outer scales L_0 and inner scales l_0, respectively.

It can be observed that the variation in the outer scale has almost no effect on the beam intensity. Moreover, the smaller the inner scale, the more gradually the LGSMV light intensity distribution broadens. When $\alpha = 3.1$ is the strongest turbulence, the inner scale has a greater influence on the LGSMV beam intensity distribution; when $\alpha = 3.5$, it has less influence.

Figure 48.23a, b show the relative beam width versus transmission distance curves for LGSMV beams with different orders and topological charges, respectively, after turbulent transmission. Figure 48.23a indicates that the relative beam width gradually decreases as n increases. Figure 48.23b indicates that the relative beam width increases as the topological charge number increases. Figure 48.23c indicates that the relative beam width increases more rapidly as the generalization factor increases for a larger turbulence intensity.

Figure 48.24a, b show the variation curves of the relative beam width of the LGSMV beam after non-Kolmogorov turbulence and free-space transmission at different wavelengths and coherence lengths, respectively. The shorter the wavelength and the larger the coherence length, the faster the relative beam width

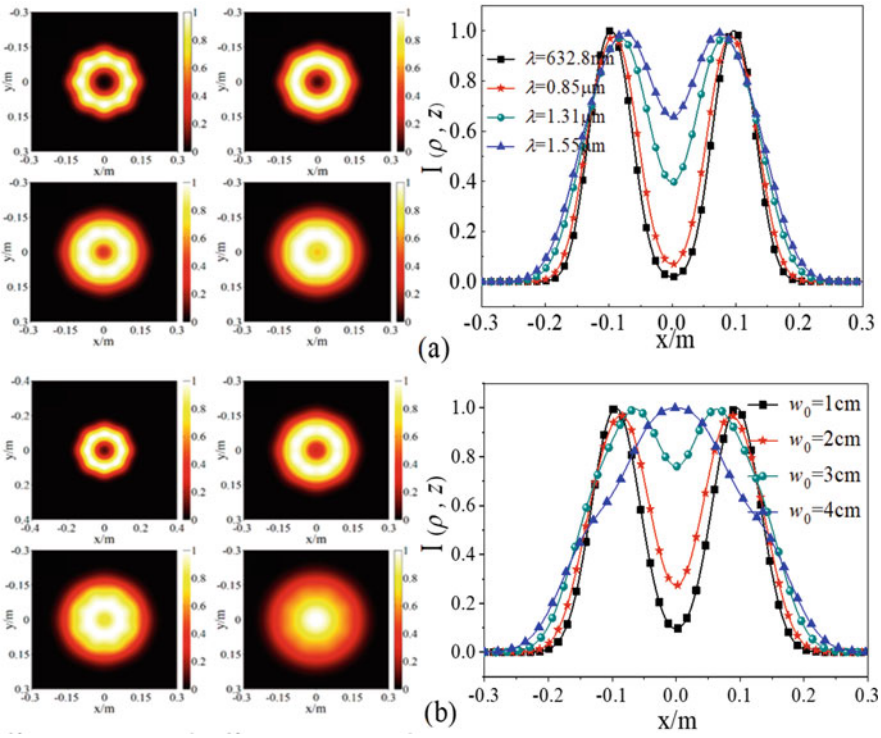


Fig. 48.21 Intensity distribution and curves of radial partially coherent vortex beam arrays transmitted through non-Kolmogorov turbulence at different parameter values. (**a**) Wavelength; (**b**) beam waist radius; (**c**) coherence length; (**d**) turbulence intensity [184]

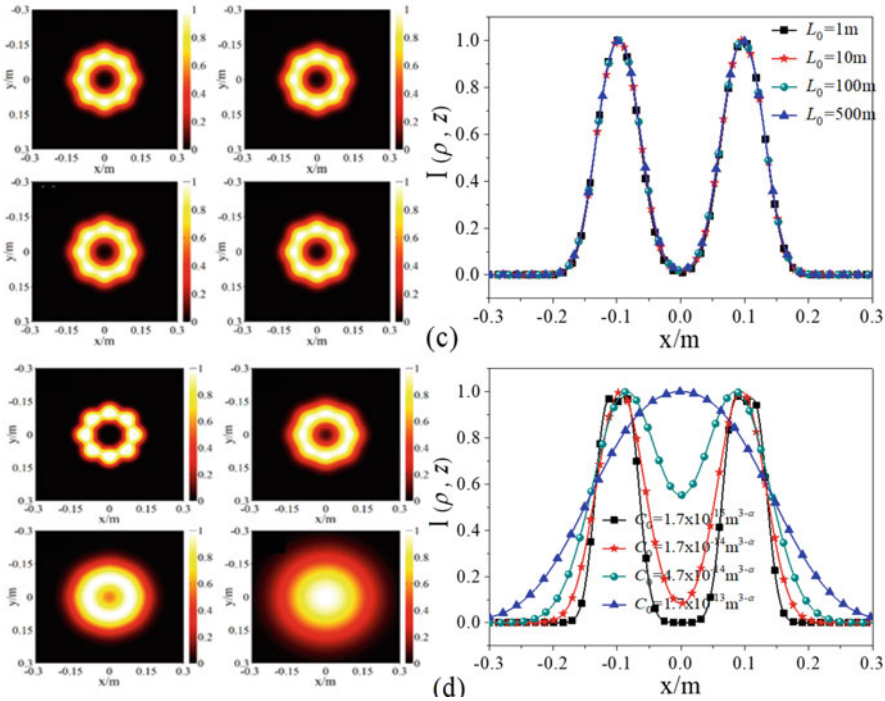


Fig. 48.21 (continued)

increases with the transmission distance, and the more serious the effect of atmospheric turbulence on the beam spreading.

An electromagnetic Gaussian–Schell vortex beam (EGSMV) was obtained by spiral phase modulating an electromagnetic Gaussian–Schell mode beam (EGSM), and the effects of the amplitude component and internal and external atmospheric turbulence scales on the degree of polarization (DoP) were analyzed.

Figure 48.25 shows the DoP distribution of the EGSMV beam at different wavelengths. It can be seen that when the wavelength is longer, the DoP distribution of the EGSMV beam after turbulence is more dispersed and the beam is more influenced by turbulence.

Figure 48.26 shows the DoP distribution curves of the EGSMV beam with $l = 1$ at different beam waist radii after a turbulent transmission. Regardless of whether the beam waist radius components w_{0x} and w_{0y} are equal, the DoP distribution of the EGSMV beam after a turbulent transmission becomes more dispersed as the beam waist radius increases.

Figure 48.27 shows the DoP distribution curves after the turbulent transmission of an EGSMV beam with $l = 1$ for different amplitude components, A_x and A_y. It can be seen that the DoP distribution of the partially coherent light source EGSMV beam after a turbulent transmission is related to the amplitude ratio A_x/A_y, not to amplitude

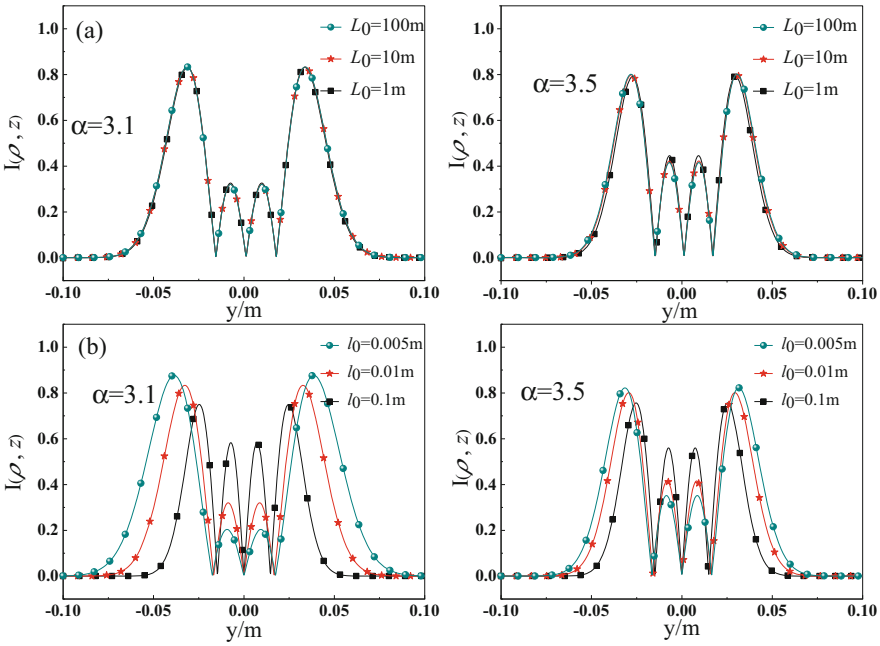


Fig. 48.22 Light intensity distribution curves of LGSMV beams transmitted by non-Kolmogorov turbulence at different inner and outer scales. (**a**) Outer scale; (**b**) inner scale [184]

A_x or A_y alone. When A_x/A_y is larger, the DoP distribution is more divergent and the beam is more influenced by turbulence.

Figure 48.28 shows the DoP distribution curves of the EGSMV beam at different inner and outer turbulence scales. For the same outer scale case, the larger the inner scale, the more concentrated the DoP, and the larger the DoP at a point on the axis, the less influenced by turbulence. The outer atmospheric turbulence scale has almost no effect on the DoP distribution.

In space quantum communication, the photons in the entangled state are not only affected by the absorption and scattering of photons by atmospheric particles but also by the decoherence of the atmospheric turbulence on the entangled photon pairs. This leads to the evolution of the entangled photons into separable states, which can degrade the performance of space quantum communication.

Figure 48.29a–d show the symbiotic entanglement and entanglement-variation curves of different OAM-entangled photon pairs under different turbulence intensities, respectively. It was found that the larger the transmission distance and turbulence intensity, the faster was the decay; the larger the OAM value, the slower was the decay. From the maximum entangled state at the beginning, that is, when the entanglement degree is 1, it gradually changes to an entangled state and finally becomes separable.

Figure 48.30 shows the joint detection probability distribution of OAM-entangled photon pairs in the atmospheric channel at different transmission distances, signal

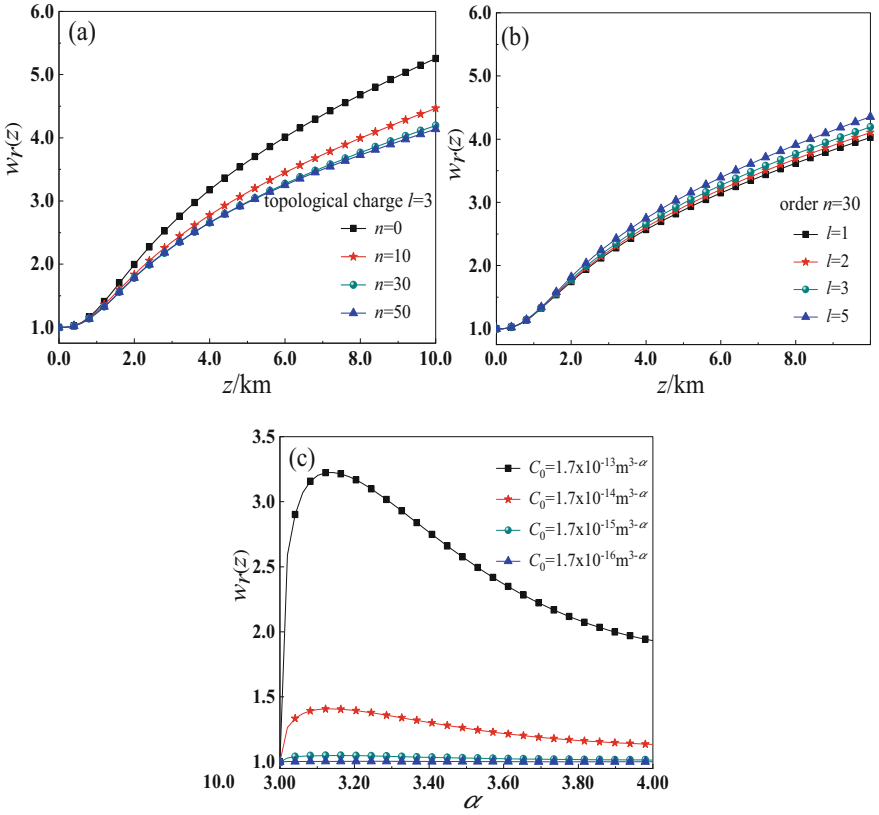


Fig. 48.23 Relative beam width of LGSMV beams after non-Kolmogorov turbulence with varying propagation distances. (**a**) Different orders n; (**b**) different topological charges l; (**c**) different turbulence intensities [184]

light wavelengths, and turbulence intensities, respectively. It can be observed that the larger the transmission distance, the stronger the turbulence intensity, and the smaller the signal light wavelength, the smaller the detection probability of the OAM state $|l_1; l_2 = l_p - l_1\rangle$. The joint detection probability of the two-photon OAM state $|l_1; l_2 \neq l_p - l_1\rangle$ then increases, and the entangled photon pair is more influenced by turbulence.

Figure 48.31 shows some transmission characteristic curves of photonic OAM. The curve in Fig. 48.31a illustrates that the nonmetallic ore affects the average per single-photon OAM to a lesser extent than metallic iron. Figure 48.31b–d show the effects of a tilt aberration, coma aberration, and astigmatism aberration on the photon OAM detection probability in turbulence, respectively. It is observed that the effect of the turbulent tilt aberration is the largest, followed by the coma aberration, and the smallest is astigmatism.

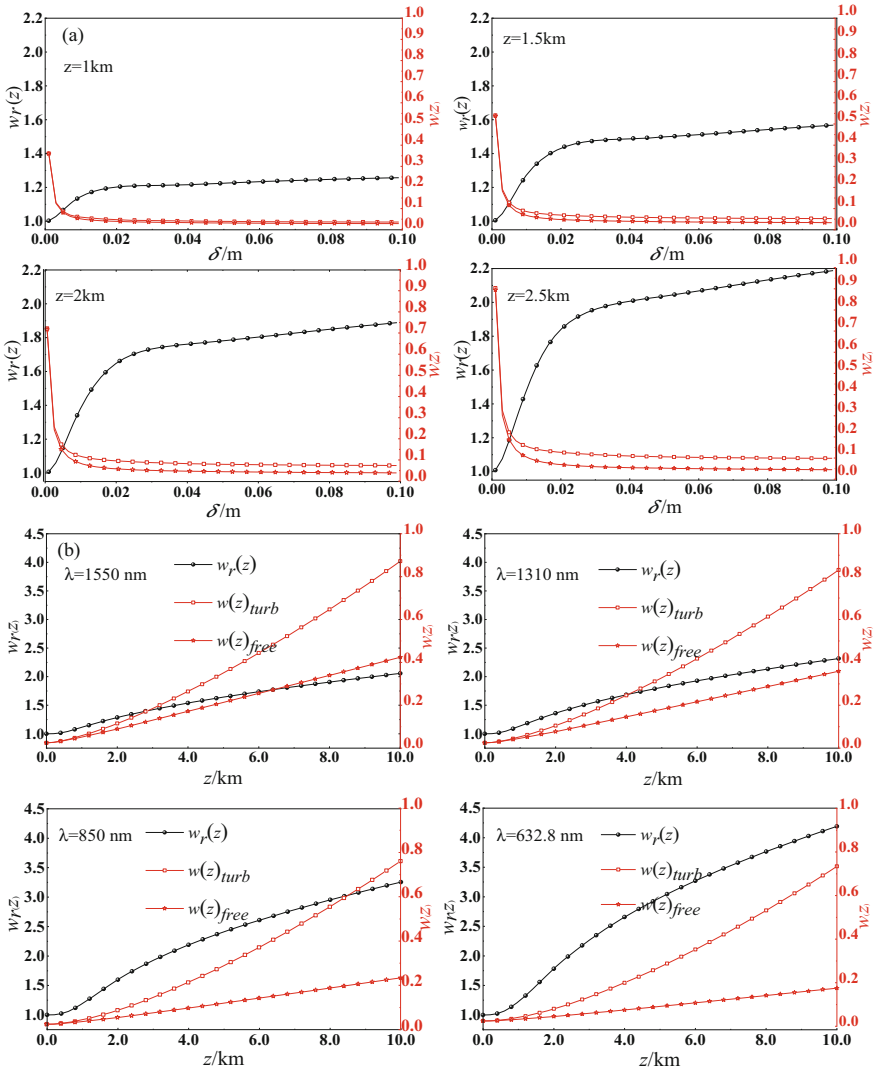


Fig. 48.24 Relative beam width curves of LGSMV beams after non-Kolmogorov turbulence with different parameters. (**a**) Wavelength; (**b**) transmission distance [184]

Finally, crosstalk suppression in an OAM-MIMO multiplexing system is studied. Figure 48.32a, b show a comparison of the constellation diagrams before and after equalization, using the cross-correlation function–multimode blind equalization (CC-MCMA) and multiuser kurtosis–multimode blind equalization (MUK-MCMA) algorithms, respectively. The MUK-MCMA algorithm restored the signal better than the CC-MCMA algorithm did.

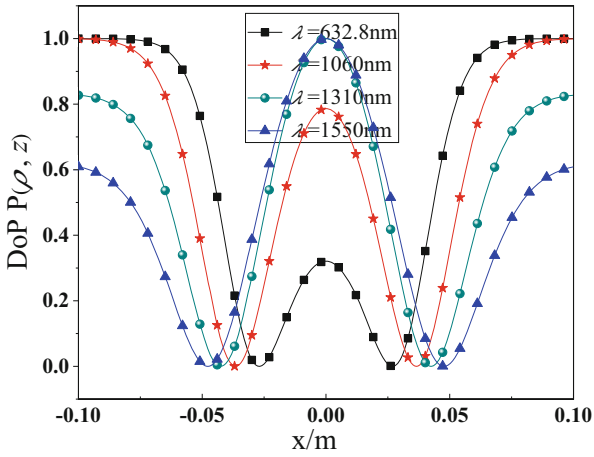


Fig. 48.25 DoP distribution curves of EGSMV beams transmitted by turbulence at different wavelengths [159]

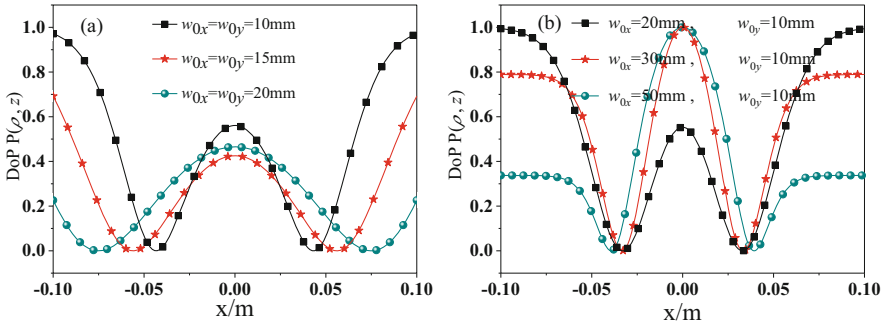


Fig. 48.26 DoP distribution curves of EGSMV beams with different waist radii after turbulent transmissions. (**a**) $w_{0x} = w_{0y}$; (**b**) $w_{0x} \neq w_{0y}$ [184]

Figure 48.33a, b show the variation curves of the average bit error rate (ABER) of the OAM-multiplexed communication system with the atmospheric turbulence intensity and the signal-to-noise ratio in the atmospheric turbulence channel, respectively. As shown in Fig. 48.33a, the MUK-MCMA algorithm can reduce the system ABER by 1–3.5 orders of magnitude and 0.4–1 orders of magnitude under either strong or weak turbulence, respectively.

In Fig. 48.33b, when the signal-to-noise ratio is approximately 20 dB, the MUK-MCMA algorithm improves the ABER equalization effect by one order of magnitude, compared with the CC-MCMA algorithm; therefore, the MUK-MCMA algorithm has the best equalization effect on the OAM multiplexing system.

Adaptive Optical Wavefront Correction

A vortex beam experiences wavefront phase distortion after atmospheric turbulence, and adaptive optical correction techniques are required to recover as much light

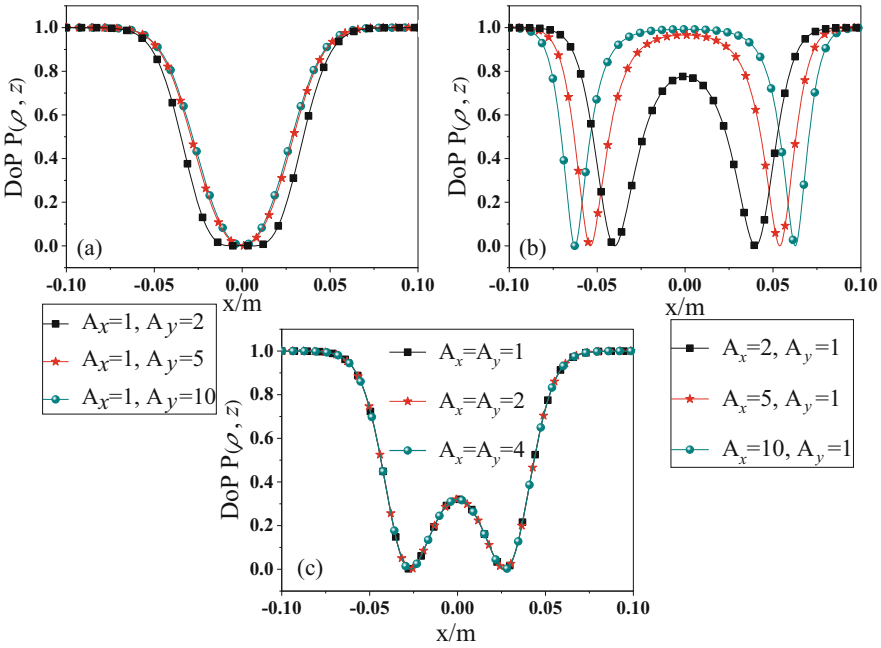


Fig. 48.27 DoP distribution curves of EGSMV beams transmitted through atmospheric turbulence with different amplitudes. (**a**) $Ax = Ay$; (**b**) $Ax < Ay$; (**c**) $Ax > Ay$ [184]

intensity phase information of the beam as possible. Here, adaptive optical correction techniques for wavefront-free measurements, based on the GS phase-retrieval algorithm, SPGD algorithm, and PD algorithm, are studied and analyzed.

The GS algorithm recovers the desired phase information by collecting the light intensity information in both the input and output planes, iterating cyclically under certain conditions to obtain the distorted phase information, and then loading the phase grayscale diagram onto a corrector, such as an SLM, by taking the conjugate.

Figure 48.34 shows the light intensity distribution of the single-mode and multimode vortex beams before and after correction by the GS algorithm. Figure 48.34 (a1)–(d1) are the spots without turbulence, Fig. 48.34(a2)–(d2) are the distorted spots after turbulence, and the corrected spots are Fig. 48.34(a3)–(d3). This shows that the singularity of the beam no longer wanders and is more uniformly distributed. The petal-like spot in the multimode case is also more complete.

The SPGD algorithm calculates the optimal voltage value of the deformation mirror corresponding to the maximum value of the light intensity correlation coefficient, based on the change in the light intensity correlation function in the iterative algorithm, which, in turn, drives a new deformation of the deformation-mirror surface. Thus, the beam reflected from the deformation mirror changes the optical-path difference, thus correcting the wavefront distortion [189].

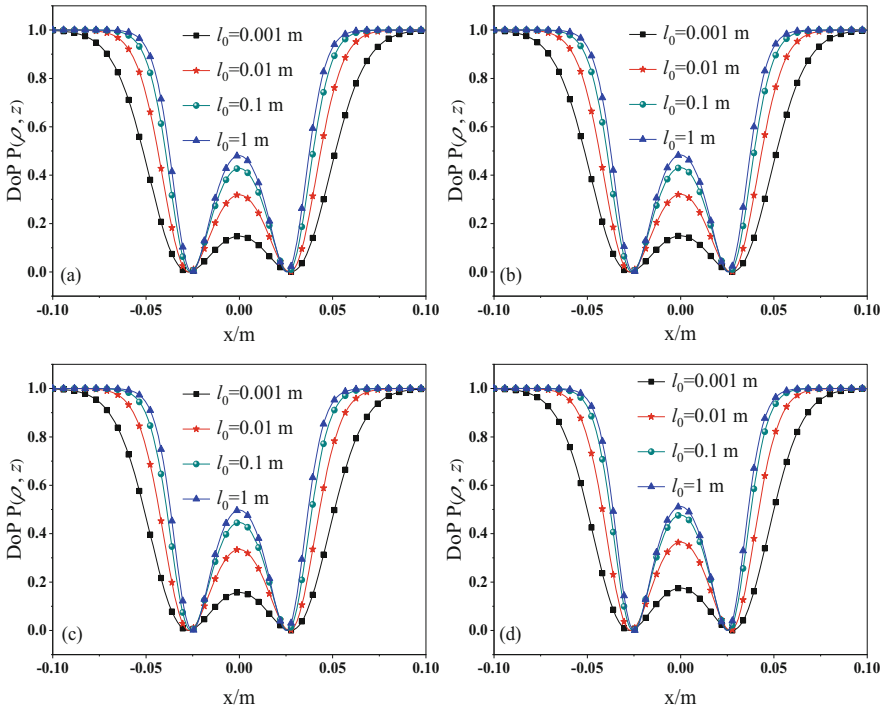


Fig. 48.28 DoP distribution curve of EGSMV beams with $l = 1$ transmitted through turbulence at different internal and external scales. (**a**) $L_0 = 200$ m; (**b**) $L_0 = 100$ m; (**c**) $L_0 = 10$ m; (**d**) $L_0 = 1$ m [184]

Figure 48.35 shows the light intensity distribution of single-mode and multimode vortex beams before and after correction using the SPGD algorithm. Figure 48.35 (a1)–(d1) show the light spots without turbulence; Fig. 48.35(a2)–(d2) show the distorted light spots after turbulence; and Fig. 48.35(a3)–(d3) show the corrected light spots. It can be seen that after correction by the SPGD algorithm, the vortex light intensity distribution becomes uniform and the spot shape becomes more regular. The correction effect is better for single-mode vortex beams than multimode beams.

The phase difference method is a phase recovery and image restoration technique. The algorithm collects the beam or image information in the focal and defocus planes, solves the optimized objective function using maximum-likelihood estimation, and obtains the phase distortion error by taking the difference between the obtained distorted phase and the ideal vortex phase. It finally recovers the original wavefront phase or image using the conjugate gray diagram of the phase distortion error.

Figure 48.36(a1)–(c1) show the beam distribution without turbulence; Fig. 48.36 (a2)–(c2) and (a3)–(c3) are the focal and defocused distorted spots after turbulence; and the corrected spots are Fig. 48.36(a4)–(c4). It can be seen that the light intensity

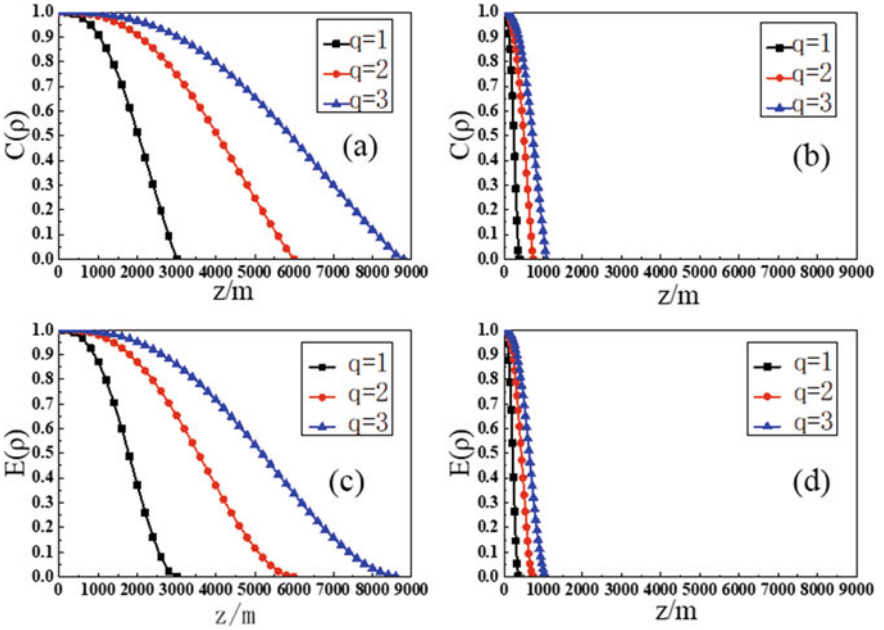


Fig. 48.29 Two-photon variation curves of different OAM entanglements. (**a–b**) Concurrence entanglement degree; (**c–d**) entanglement degrees [185]

distribution of Fig. 48.36(a2–c2) and (a3–c3) has decayed and distorted, and after correction by the phase difference method, the energy is recovered and the degree of distortion becomes lighter (Table 48.6).

Others

In addition to the above studies on transmission characteristics, there are also analyses of OAM beams passing through single-sphere particle scattering and through optical aperture systems. Figure 48.37 shows a model of OAM beam single-sphere particle scattering and the variation curve. The total scattering intensity is weaker than that of plane waves, as seen in Fig. 48.37a, and the scattering cross section increases as the beam width increases, as seen in Fig. 48.37b, c. The scattering cross section of particles is larger for plane waves than for Gaussian waves.

Figure 48.38a, b show the physical model of the LG beam passing through two aperture systems. Figure 48.39 shows the results after passing through the two optical systems. It can be found that the system obscuration ratio affects the beam intensity but does not affect the beam width, and the larger the wavelength, the greater the effect of the obscuration ratio on the beam transmission. In addition, in the near-field receiving range, as the spacing between the lens and aperture increases, the effect of the aperture and beam obscuration ratio on the beam transmission increases.

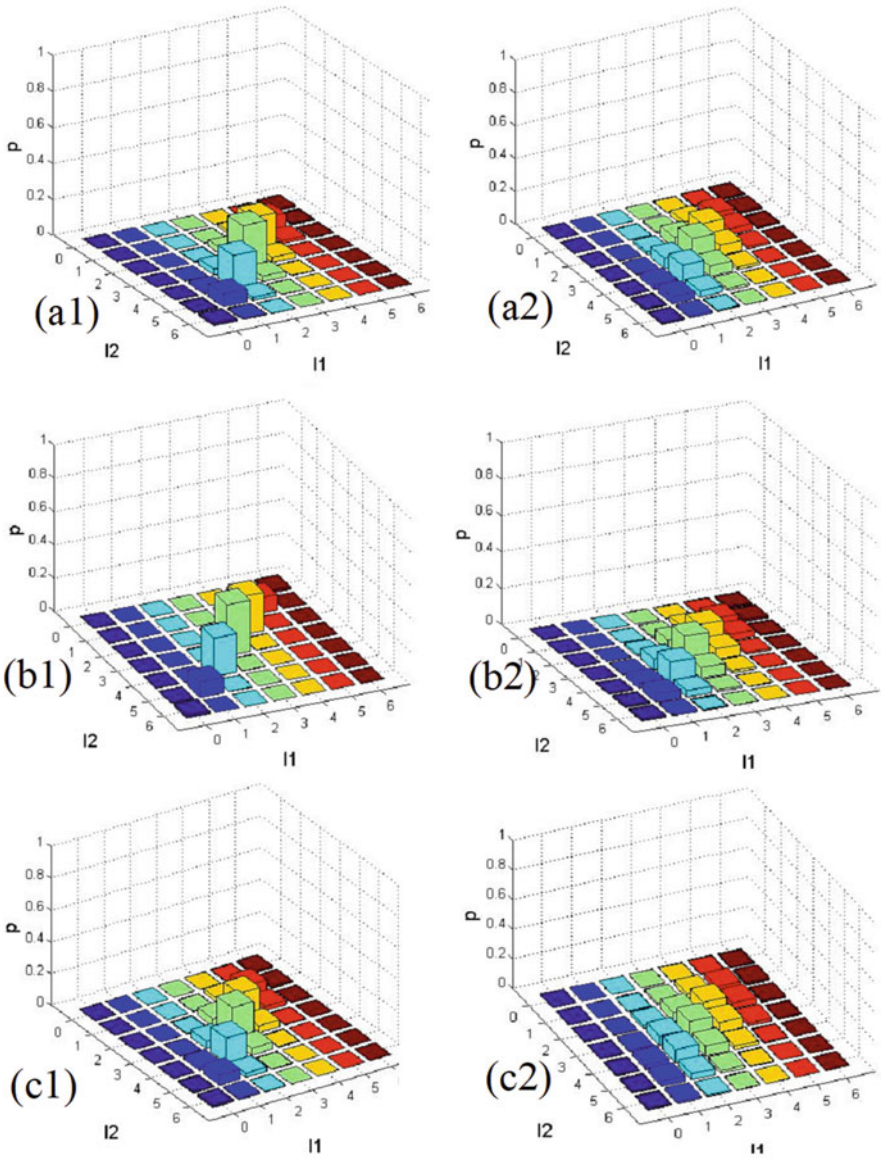


Fig. 48.30 Joint detection probability of OAM-entangled photon pairs under different transmission distances, turbulence intensities, and signal light wavelengths in atmospheric channels. (a1–a2) 100 m and 300 m; (b1–b2) $C_n^2 = 6.4 \times 10^{-17} m^{-2/3}$ and $C_n^2 = 8.7 \times 10^{-15} m^{-2/3}$; (c1–c2) 1028 nm and 632 nm [186]

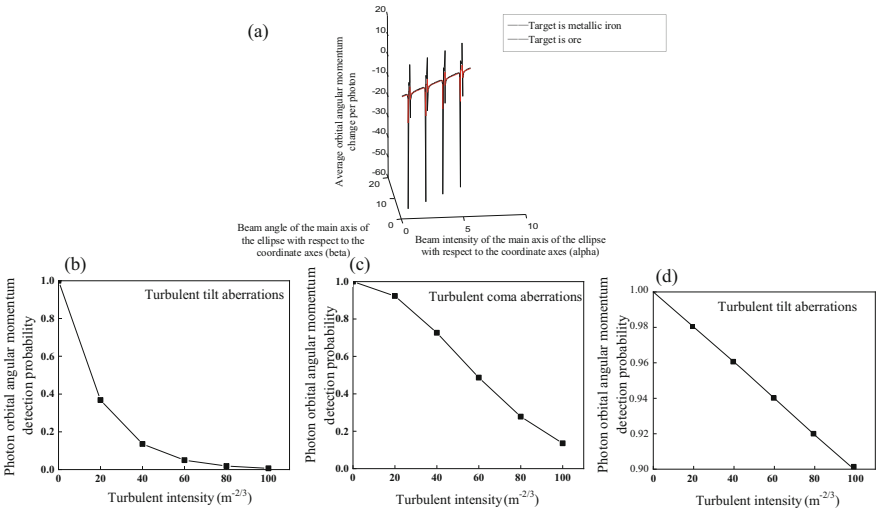


Fig. 48.31 Transmission characteristics of photon OAM. (**a**) Single-photon average OAM variation after interaction of high-order elliptic Hermite–Gaussian beams with metallic iron and ore; (**b–d**) Effects of tilt aberration, coma aberration, and astigmatic aberration on photon OAM detection probability in turbulence [187]

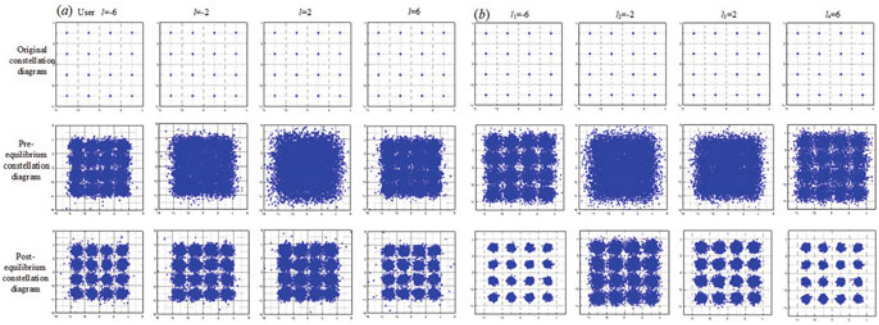


Fig. 48.32 Constellation diagram before and after equilibrium using different algorithms. (**a**) CC-MCMA; (**b**) MUK-MCMA [188]

Vortex Beam Separation and Detection

Vortex Beam Detection

At present, there are several methods for detecting OAM modes:

1. Vortex beam and plane wave to detect spherical wave interference
2. Vortex beam diffraction after passing through a slit or type of aperture

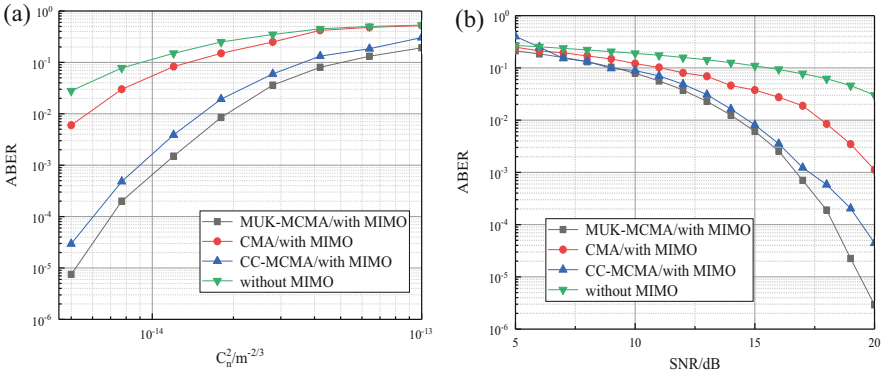


Fig. 48.33 Variation curves of the average bit error rate of an OAM-multiplexed communication system in an atmospheric turbulence channel. (**a**) Turbulence intensity; (**b**) signal-to-noise ratio (SNR) [188]

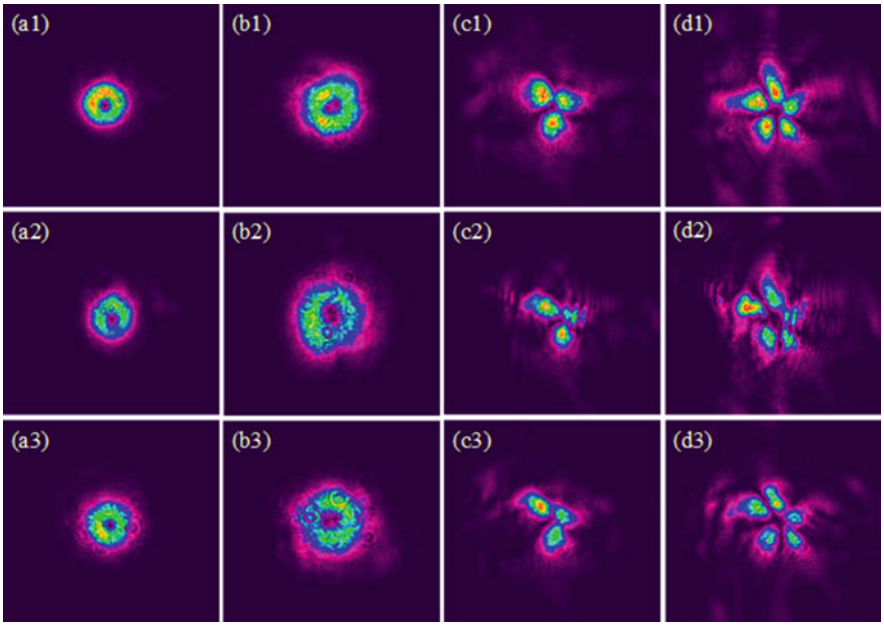


Fig. 48.34 Light intensity distribution before and after GS algorithm correction. (a1)–(d1) without turbulence; (a2)–(d2) after turbulence; (a3)–(d3) after correction [189]

3. Vortex beam after diffraction by optical devices, such as gratings, to reconstruct the wavefront for better detection
4. Other emerging methods, such as polarization-direction angle detection

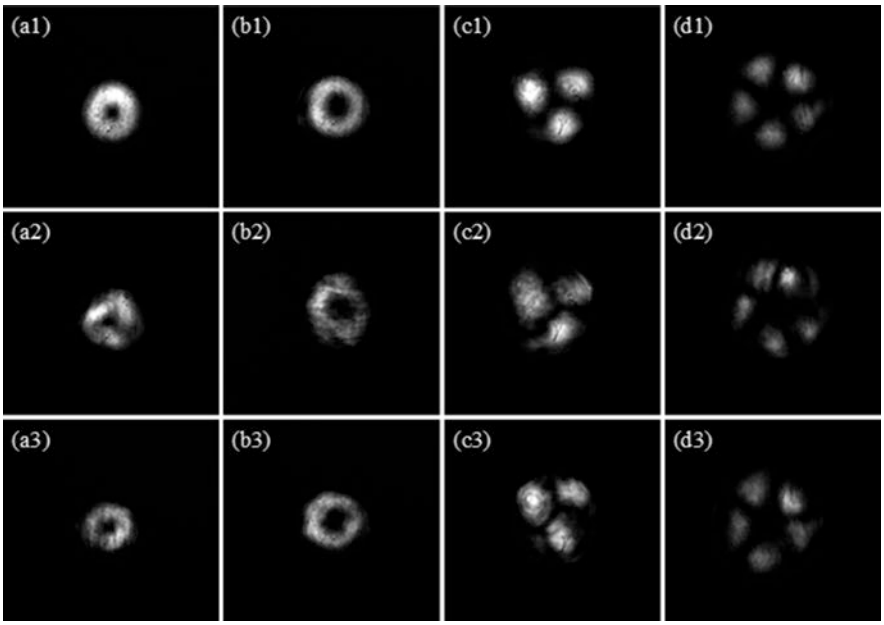


Fig. 48.35 Light intensity distribution before and after SPGD algorithm correction. (a1)–(d1) no turbulence; (a2)–(d2) after turbulence; (a3)–(d3) after correction [189]

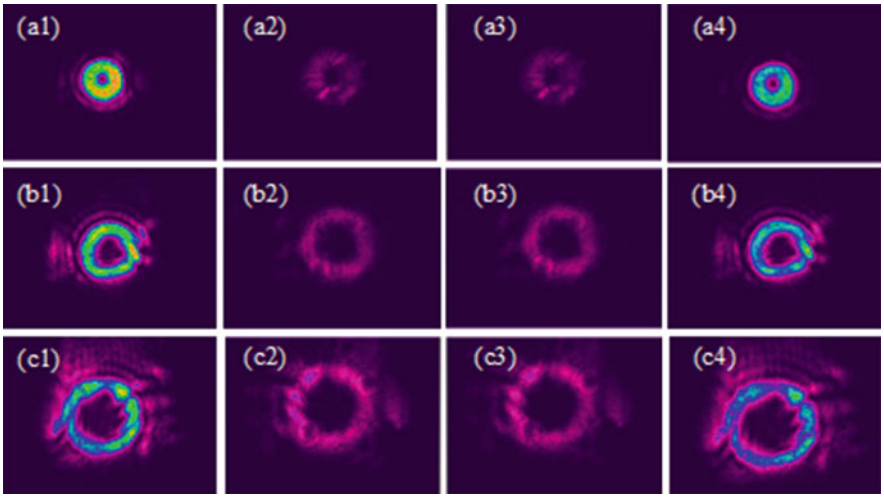


Fig. 48.36 Intensity distribution and spiral spectrum of vortex beams corrected by the phase difference method. (a1)–(c1) Initial plane; (a2)–(c2) focal plane; (a3)–(c3) defocused plane; (a4)–(c4) after correction [187, 190]

Table 48.6 Summary of wavefront sensorless correction algorithms in adaptive optics

Adaptive optical wavefront correction algorithm	Algorithm principle	Effect of correction
GS algorithm	Wavefront intensity reconstructs the wavefront phase	Single-mode and multimode can be corrected, and the energy distribution becomes uniform
SPGD algorithm	The maximum value of the light intensity correlation coefficient is calculated corresponding to the optimal driving voltage of the deforming mirror	The light intensity in the center of the spot obviously increases and the energy converges
PD algorithm	Taking the difference between the distorted phase and the ideal vortex phase to obtain the conjugate	Single-mode and multimode can be corrected for more uniform light intensity distribution and energy enhancement

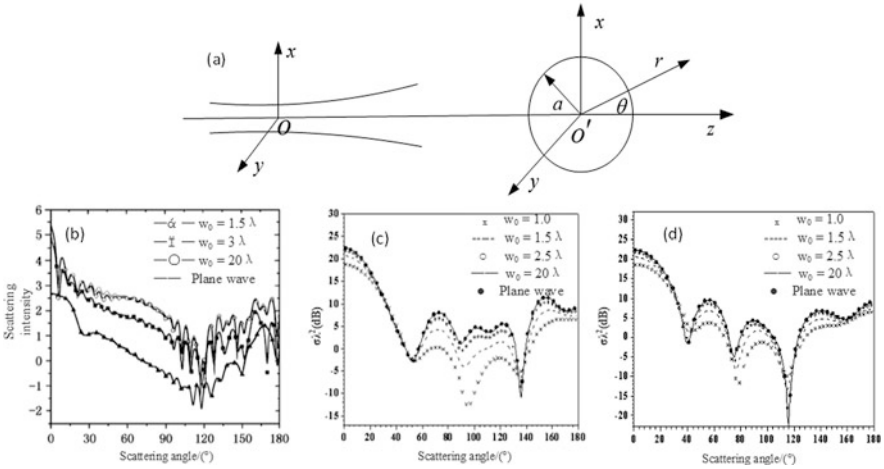


Fig. 48.37 OAM beam scattering by a single spherical particle (**a**) scattering by a single spherical particle to an on-axis Gaussian beam; (**b**) variation of scattering intensity and angular distribution with beam width; (**c**) and (**d**) effect of beam width on the scattering cross section in the E and H planes, respectively [142, 191]

Vortex Beam Interference Detection Method

Vortex beam interference detection involves using vortex light and plane light for interference superposition by analyzing the generated interference fringes to detect the topological charge of the vortex light. Figure 48.40 shows that the number of bifurcations is 1, regardless of the simulation or experiment, and the bifurcations face upward at $l = 1$ and downward at $l = -1$.

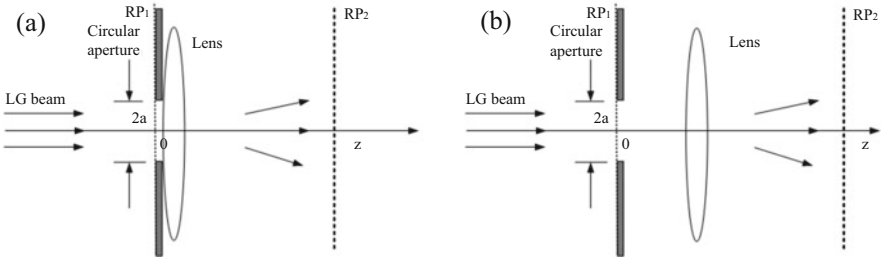


Fig. 48.38 Physical model of an optical system with a single aperture. (**a**) Lens adjacent to the aperture; (**b**) lens separated from diaphragm [142, 191]

Aperture-Detection Method

Figure 48.41 shows the diagrams of circular, square, and triangular apertures, as well as single and double slits. After passing through these apertures, the diagrams of the vortex beam can detect the OAM mode.

As shown in Fig. 48.42, the experimental results for a vortex beam diffracted by different apertures are shown. The positive and negative topological charges can be determined from the twisting direction of the diffracted spots, whereas the degree of twisting, number of array spots, and number of stripes can determine the size of the topological charge.

Table 48.7 presents a comparison of the detection effects of the five diffractive optical elements. Except for the circular aperture, which can only detect the topological charge of the incident vortex beam, the square, triangular, single, and double slits can detect the topological charge of the incident vortex beam, as well as the sign.

Grating Detection Method

The diffracted optical field of the vortex beam after passing through the gradually changing-period grating and annular grating can also be detected in the OAM mode. Figure 48.43 shows the ab-type gradually changing-period amplitude grating, ab-type gradually changing-period phase grating, annular amplitude, annular phase grating, phase correction, and fan-out phase diagram.

The light field distribution of various grating diffractions, improved by phase-correction and fan-out phase diagrams, were obtained according to the calculation and analysis of the Fresnel diffraction theory, as shown in Fig. 48.44. The topological charge magnitude and sign of the incident vortex beam can be detected by determining the number of diffraction stripes and their orientations.

The diffracted light intensity of the annular phase grating was the best, with clear stripes and the largest light intensity value. After phase correction, the maximum detectable topological charge was up to ± 20 orders for the gradually changing-period grating (GCPG) and up to ± 30 orders for the annular grating (AG).

Phase Singularity Evolution

Figure 48.45 shows the phase singularity evolution in the far field after the transmission of LGSMV beams with different topological charges and orders. With an

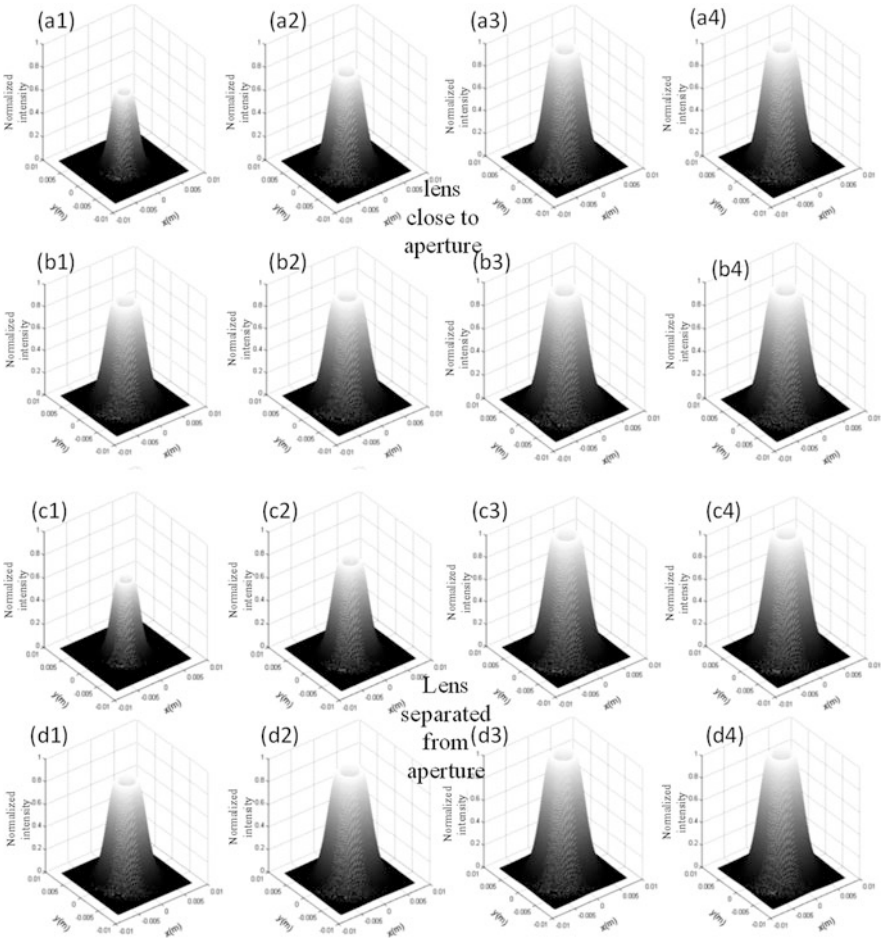


Fig. 48.39 Light intensity distribution of LG_0^1 under different obscuration ratios a/w_0 and wavelengths. (a1–a4) Obscuration ratio for $\lambda = 632.8$nm is $a/w_0 = 1.2, 1.6, 1.8, 2.0$; (b1–b4) obscuration ratio for $\lambda = 1060$nm is $a/w_0 = 1.6, 1.8, 2.6, 3.5$; (c1–c4) obscuration ratio for $\lambda = 632.8$nm is $a/w_0 = 1.6, 1.8, 2.2, 3.5$; (d1–d4) obscuration ratio for $\lambda = 1060$nm is $a/w_0 = 1.5, 2, 3.4, 6$ [142, 191]

increase in the transmission distance, the LGSMV light intensity distribution gradually expands, the central dark region splits, and the total number of phase singularities is equal to the original topological charge.

Orientation Angle of Polarization

The orientation angle of polarization (OAoP), which characterizes the direction of polarization of a vector vortex beam, can also be used to detect the topological

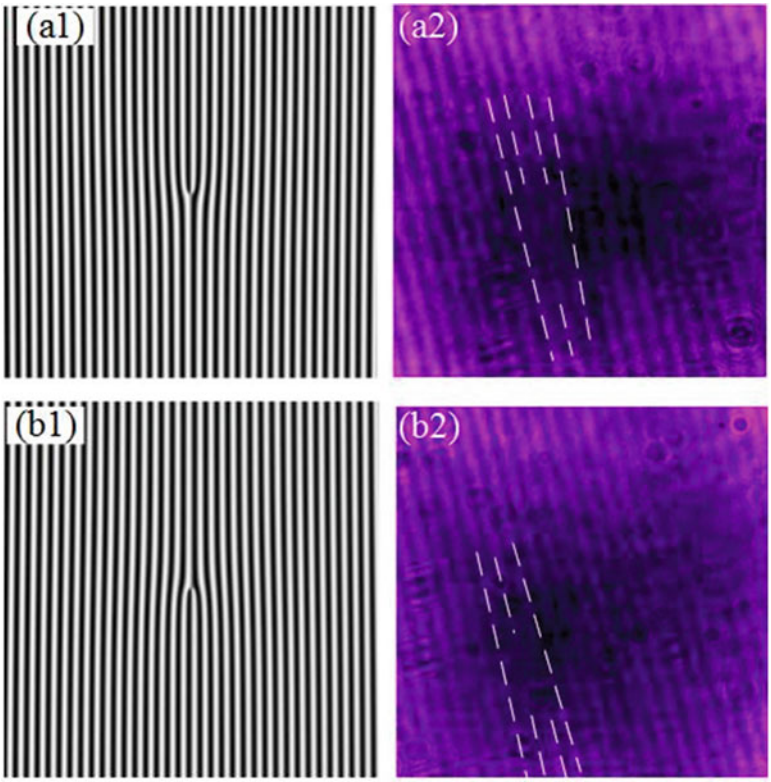


Fig. 48.40 Interference patterns of two vortex beams. (a1–a2) Vortex interference pattern with $l = 1$; (b1–b2) vortex interference pattern with $l = -1$ [145]

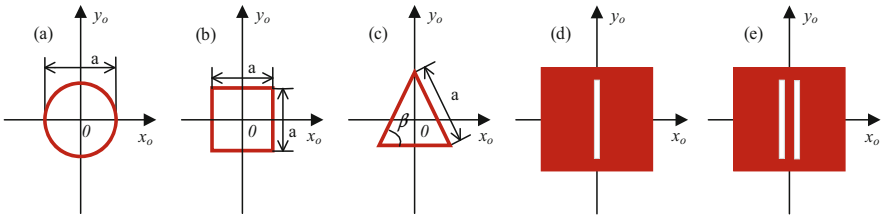


Fig. 48.41 Apertures. (**a**) Circular aperture; (**b**) square aperture; (**c**) triangular aperture; (**d**) single slit; (e) double slit [184]

charge magnitude and sign of a vortex beam. Figure 48.46 (a1–j1) are the vortex light fields with topological charge $l = \pm 1, \pm 2, \pm 3, \pm 4, \pm 5$, and (a2–j2) are the corresponding OAoP distributions. The topological charge size is equal to half the number of petals. The petals rotate counterclockwise when the topological charge is positive and vice versa when it is negative.

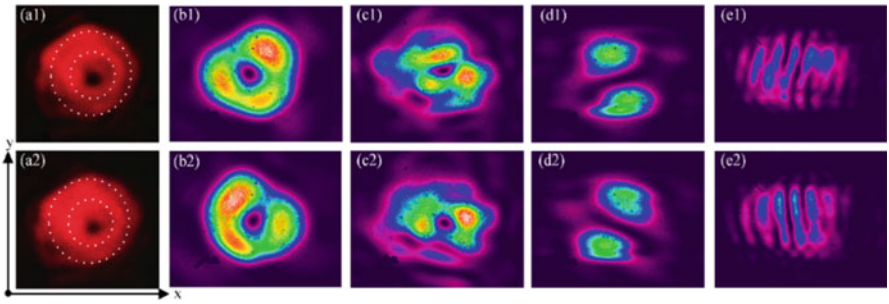


Fig. 48.42 Diffraction patterns of vortex beams with topological charges. (a1)–(e1) $l = 1$; (a2)–(e2) $l = -1$ passing through an aperture; (a1–a2) circular aperture; (b1–b2) square aperture; (c1–c2) triangular aperture; (d1–d2) single slit; (e1–e2) double slit [184]

Table 48.7 Comparison of the detection effect of five diffractive diaphragms

Diaphragm diffraction Topological load number	Round hole diaphragm	Triangular hole diaphragm	Square hole diaphragm	Single seam	Double seam
Size judgment method (number of diffraction streaks or bright spots)	$\|l\| + 1$	$(\|l\|^2 + 3\|l\|)/2 + 1$	Odd numbers: $2(\|l\| + 1)$ Even: $(\|l\|/2 + 1)^2$	$\|l\|-1$	The degree of distortion of the stripes
Symbol judgment method (light spot facing or twisting direction)	/	Left positive and right negative	Left positive and right negative	Left positive and right negative	Left positive and right negative

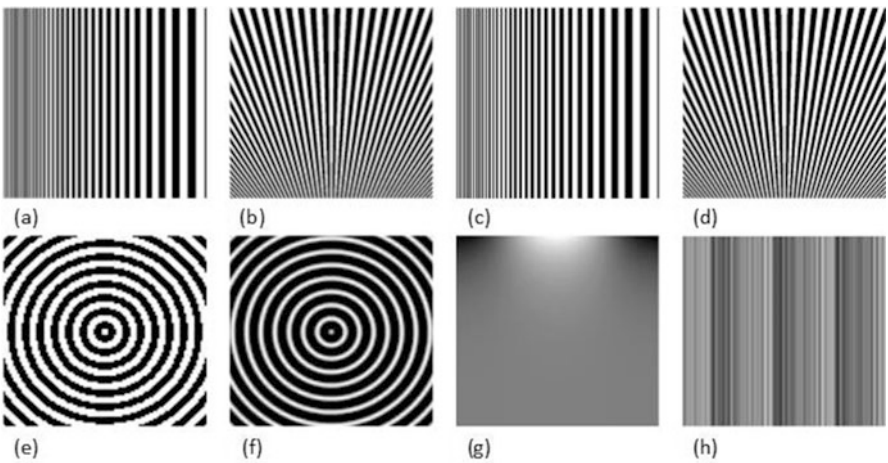


Fig. 48.43 Measurement gratings and phase diagrams [192]

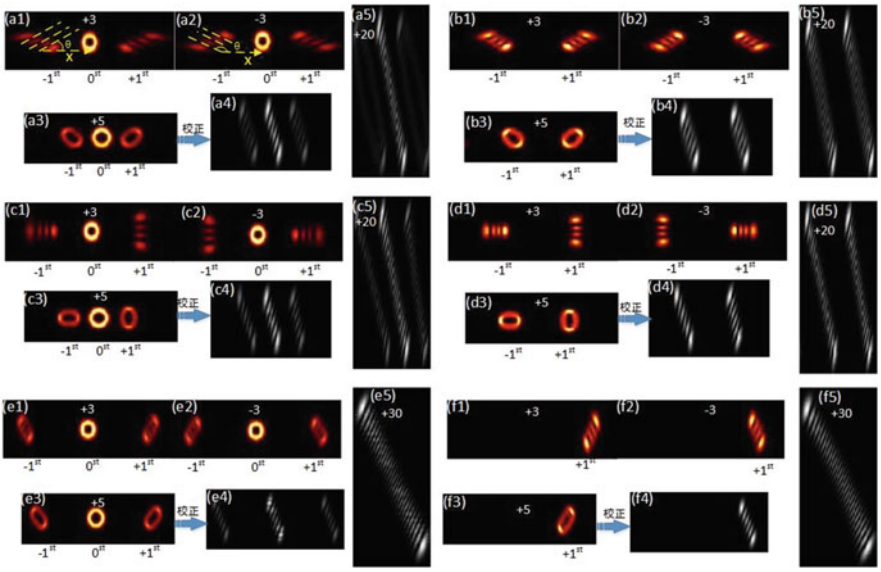


Fig. 48.44 Measurement results of the gradually changing-period grating and annular grating. (a1–a5) a-type gradually changing-period amplitude grating; (b1–b5) a-type gradually changing-period phase grating; (c1–c5) b-type gradually changing-period amplitude grating; (d1–d5) b-type gradually changing-period phase grating; (e1–e5) annular amplitude grating; (f1–f5) annular phase grating [192]

Vortex Beam OAM Separation and Detection

Figure 48.47 shows the diffracted light field distributions of $l = 1, 3$, $l = 1, 5$, and $l = 1, 7$ after a coordinate transformation. The coordinate-transformed light spot is horizontally laterally distributed in the form of bright stripes (Table 48.8).

Vortex Beam OAM Applications

Vortex Beam OAM Coding and Decoding

In OAM multiplexing communication, the additional degrees of freedom of the vortex beam can be used to realize multibit data coding and decoding for higher speed and capacity communication. Here, the phase coding is based on the eight-stage structure of an SPP, that is, each eight-stage structure can encode 24 bits of information after encoding by equal length, and the value and ordering of these eight characters represent phase coding.

The spiral spectrum is measured, and the value of the spiral spectrum is stored with the encoded information of the QSS through the correspondence of the step encoding and spiral spectrum to decode the information. The spiral spectrum has symmetric and asymmetric structures, as shown in Fig. 48.48a, b, respectively.

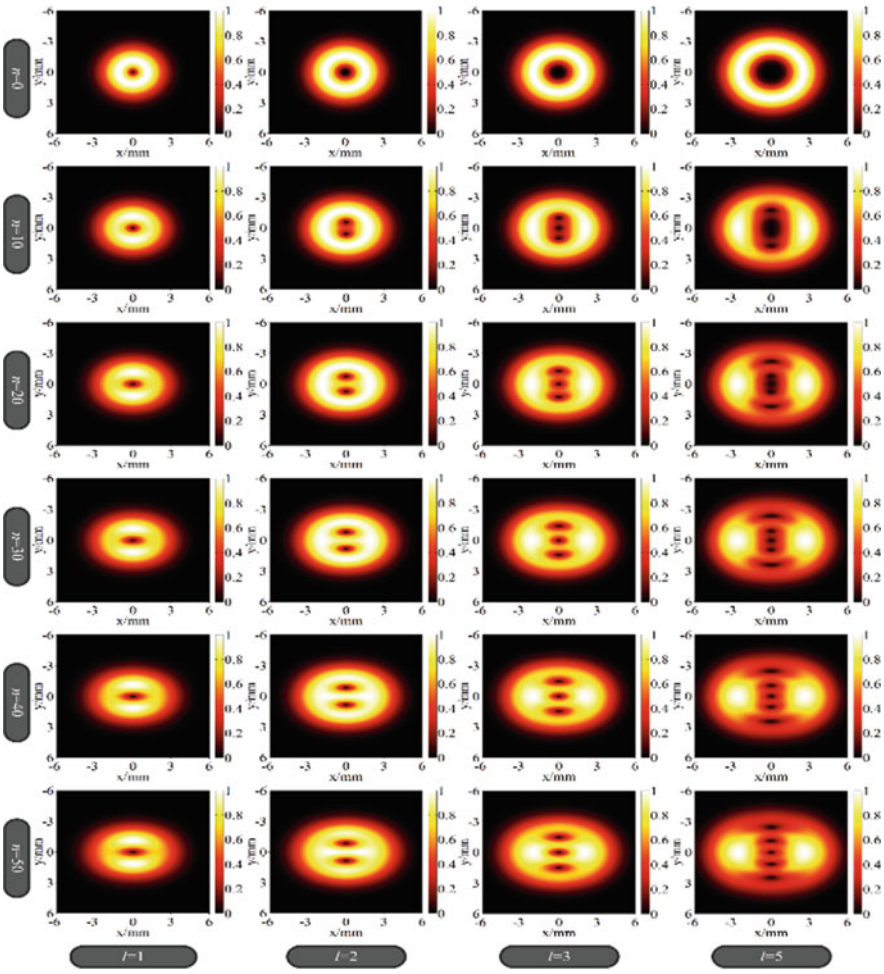


Fig. 48.45 Phase singularity evolution of LGSMV beams with different topological charges and orders [184]

The conical lens is used as a mode converter to convert the LG beam into a higher-order Bessel diffraction-free beam for spatial light information-transmission encoding, and then the array grating is used for decoding. Figure 48.49 shows the schematic diagram of the coding and decoding system.

The optical communication implementation method is based on orthogonal OAM state information encoding using a high-order Bessel function. A single photon can carry a determined OAM state, which can increase the key transmission rate of the protocol and provide a solution to simultaneously increase the optical communication distance. Figure 48.50 shows the high-order Bessel beam optical quantum communication codec system.

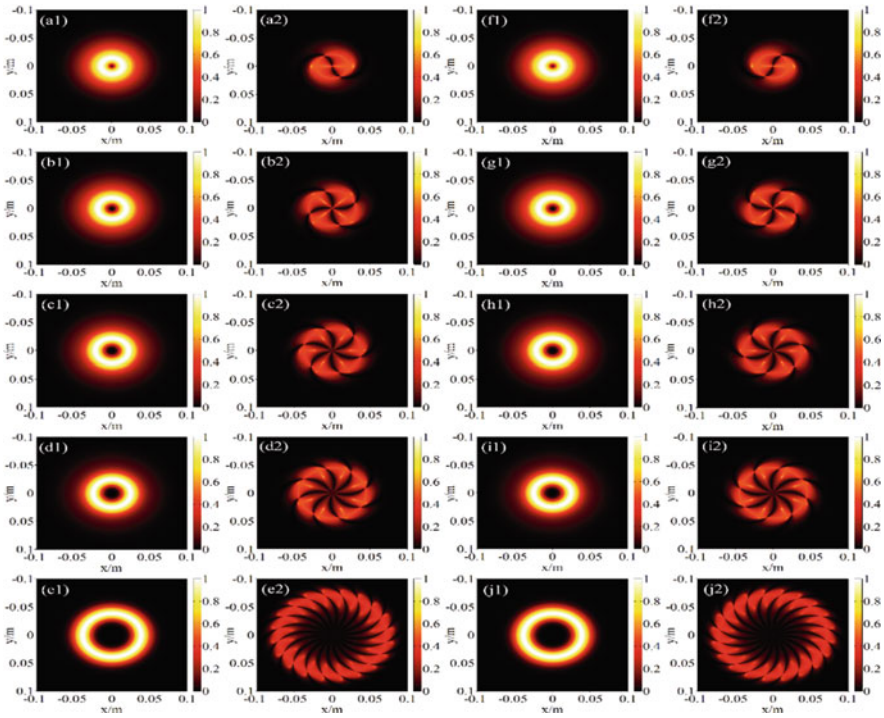


Fig. 48.46 Far-field diffraction patterns of EGSMV beams with different topological charges [184]

Vortex Beam OAM Information Exchange

OAM multiplexing communication systems exchange different data streams with different OAM modes. The topological load conversion of different OAM beams can increase the data processing flexibility and management in the communication link and realize a reconfigurable optical network based on OAM.

Figure 48.51 shows a schematic diagram of the OAM beam information exchange principle, which enables the OAM mode to exchange information by loading the spiral phase diagram on the SLM. Figure 48.52 shows the experimental results of different OAM beam information exchanges at different carrier frequencies. It can be clearly seen that the signal waveforms at different frequencies are switched, and the information of the two OAM beams is exchanged.

Vortex Beam OAM Through Maca-Antenna Emission-Efficiency Analysis

Maksutov–Cassegrain antennas (Maca-antennas) have the advantages of lower spherical aberrations, chromatic aberrations, and a long focal length, which can improve the resolution of the antenna system. Therefore, they are widely used in satellite remote sensing and LIDAR systems. Here, it is proposed that using a spirally distributed

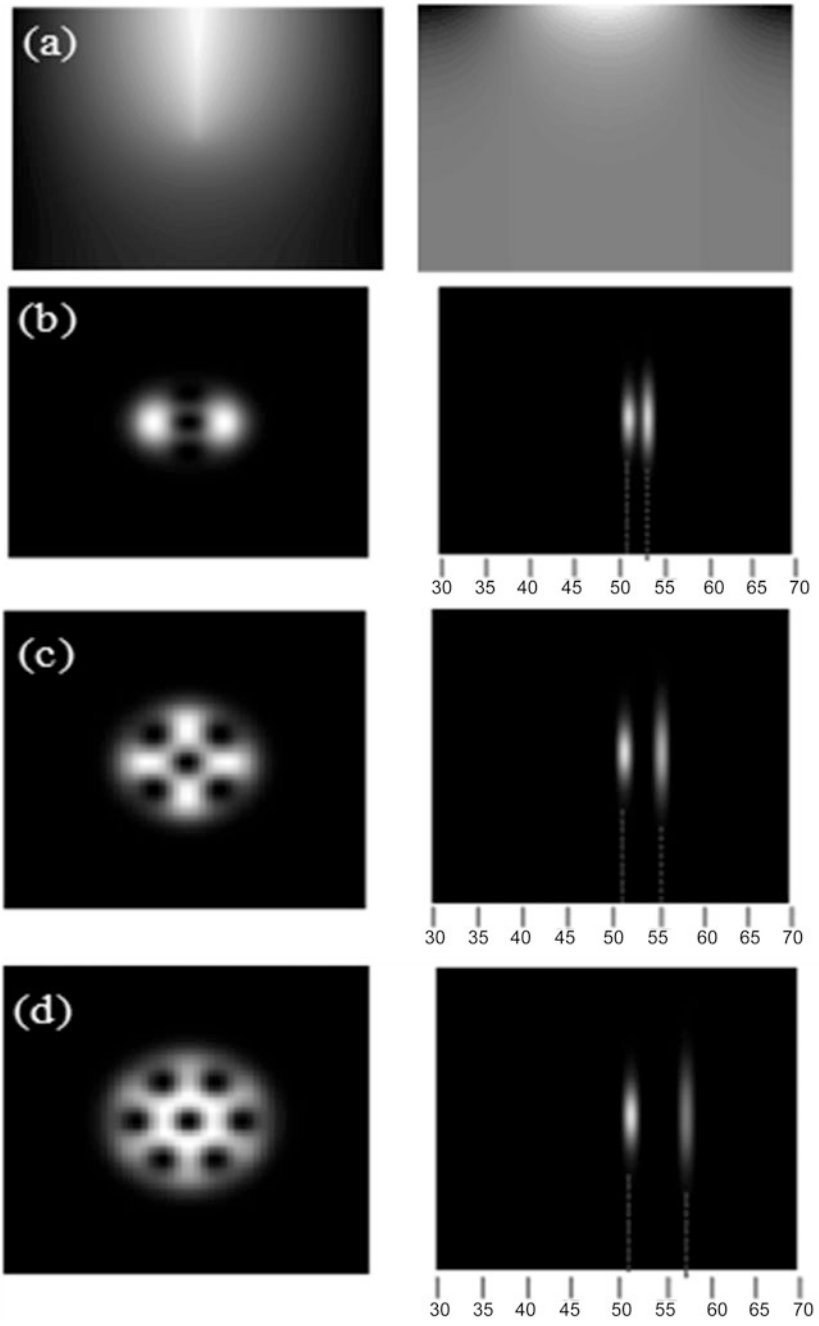


Fig. 48.47 Light field superimposed by different topological charges and light field after coordinate transformation. (**a**) Phase mask of coordinate transformation; (**b**) $l = 1, 3$; (c) $l = 1, 5$; (d) $l = 1, 7$ [178]

Table 48.8 Comparison of diffraction measurement effects

Diffraction topology	Magnitude determination method (diffracted spot or stripes)	Sign determination method (spot orientation or twist direction)	Measurement performance
Interference method	Stripe bifurcation number	Up: positive Down: negative	Low order/ topological charge less than 5
Circular aperture	$\|l\| + 1$	–	Low order/ topological charge less than 5
Triangular aperture	$(\|l\|^2 + 3\|l\|)/2 + 1$	Left: positive Right: negative	Low order/ topological charge less than 10
Square aperture	Odd: $2(\|l\| + 1)$ Even: $(\|l\|/2 + 1)^2$	Left: positive Right: negative	Low order/ topological charge less than 10
Single slit	$\|l\| - 1$	Left: positive Right: negative	Low order/ topological charge less than 5
Double slit	Degree of fringe distortion	Left: positive Right: negative	Low order/ topological charge less than 5
A-type gradually changing-period amplitude/phase grating	$\|l\| - 1$	First-order diffraction order with x-axis positive angle: $\theta \leq 90^\circ$is positive; $90^\circ < \theta \leq 180^\circ$ is negative	Topological charge less than 10/good diffraction effect for phase grating
B-type gradually changing-period amplitude/phase grating	$\|l\| - 1$	First-order diffraction order with x-axis positive angle: $\theta \leq 90^\circ$is positive; $90^\circ < \theta \leq 180^\circ$ is negative	Topological charge less than 10/good diffraction effect for phase grating
Annular amplitude/phase grating	$\|l\| - 1$	First-order diffraction order with x-axis positive angle: $\theta \leq 90^\circ$is positive; $90^\circ < \theta \leq 180^\circ$ is negative	Topological charge less than 25/good diffraction effect for phase grating
Phase correction/ fan-out	$\|l\| - 1$	First-order diffraction order with x-axis positive angle: $\theta \leq 90^\circ$is positive; $90^\circ < \theta \leq 180^\circ$ is negative	Boost to 20 and 30 for GCPG and AG
Coordinate transformation method	The position of the light field after the coordinate transformation is combined with the coordinate scale distance formula		Single/double modes with low order

circular vortex beam as the light source can improve the Maca-antenna emission efficiency. Figure 48.53 shows a schematic diagram of the LG beam passing through the Maca-antenna system, where the incident beam first strikes the secondary reflector, then reflects to the primary reflector, and finally emits the LG beam through the bending-moon correction mirror.

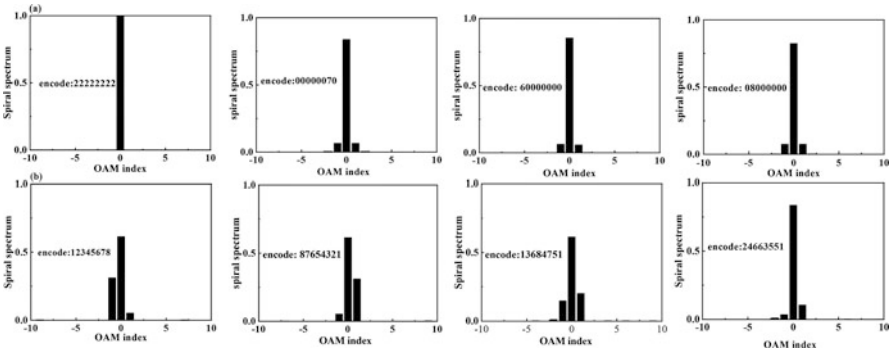


Fig. 48.48 Eight-step QSS phase coding. (**a**) Symmetric spiral spectrum; (**b**) asymmetric spiral spectrum [193]

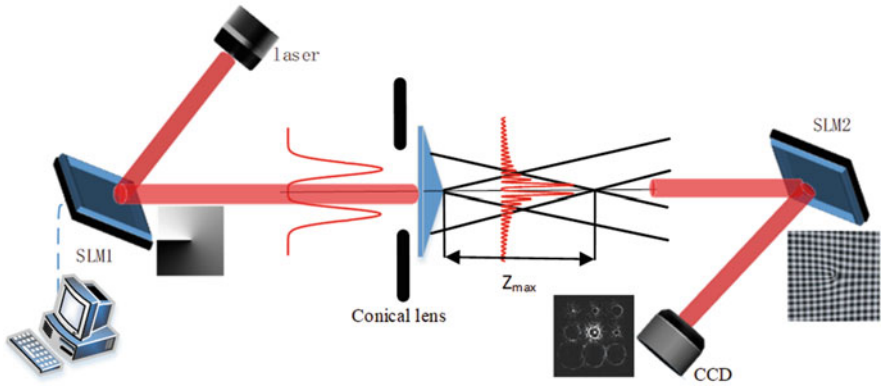


Fig. 48.49 High-order Bessel beam encoding and decoding system without diffraction [193]

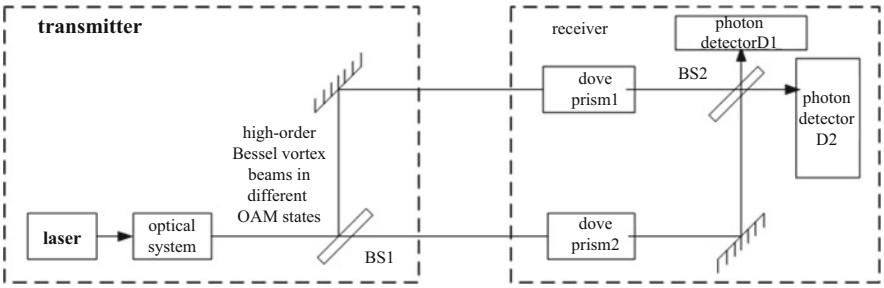


Fig. 48.50 Optical quantum encoding and decoding system of a Bessel beam OAM [193]

Figure 48.54a, b show the variation curves of the Maca-antenna transmitting efficiency with the obscuration ratio under different topological charge numbers and radial indices, respectively. Under a certain obscuration ratio, the transmitting

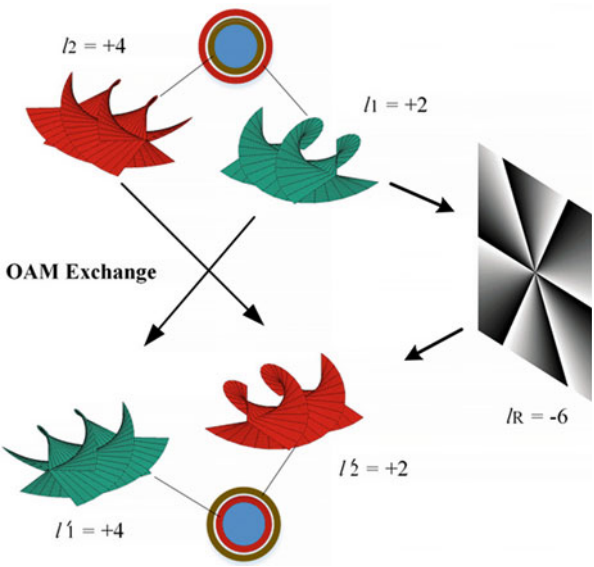


Fig. 48.51 Principle of OAM beam information transfer [194]

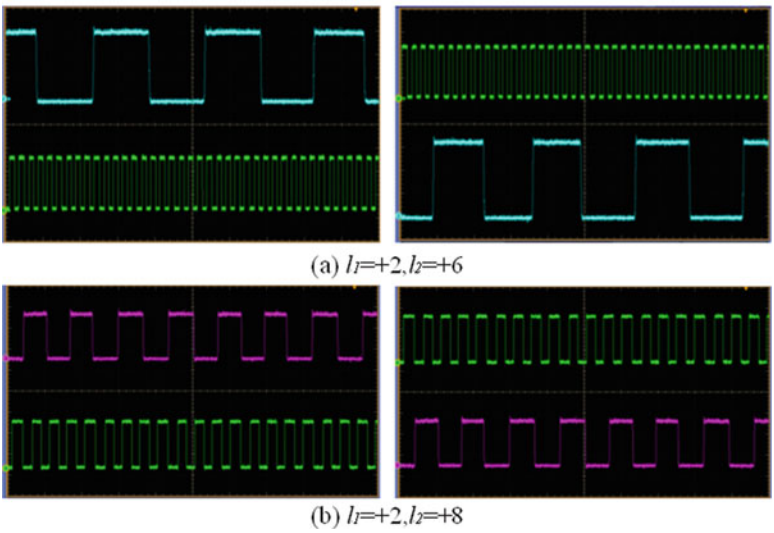


Fig. 48.52 Experimental results of OAM beam information transfer. (**a**) $l_1 = 2$ is swapped with $l_2 = 6$; (**b**) $l_1 = 2$ is swapped with $l_2 = 8$ [194]

efficiency of the vortex beam is much larger than that of a Gaussian beam, and the larger the topological charge number, the larger the transmission efficiency of the Maca-antenna; it can reach more than 80% in a wider obscuration ratio range.

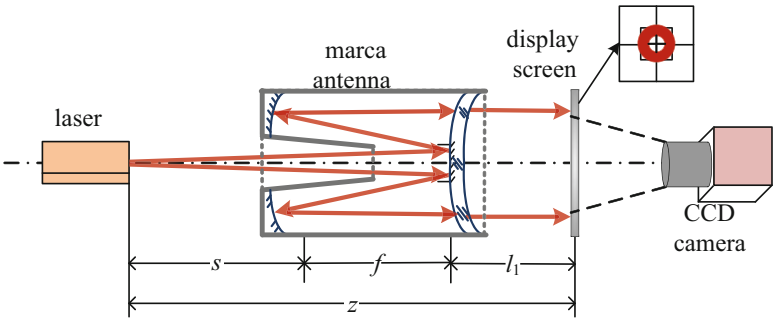


Fig. 48.53 Schematic diagram of the off-axis emission of an LG beam through a Maca-antenna [184]

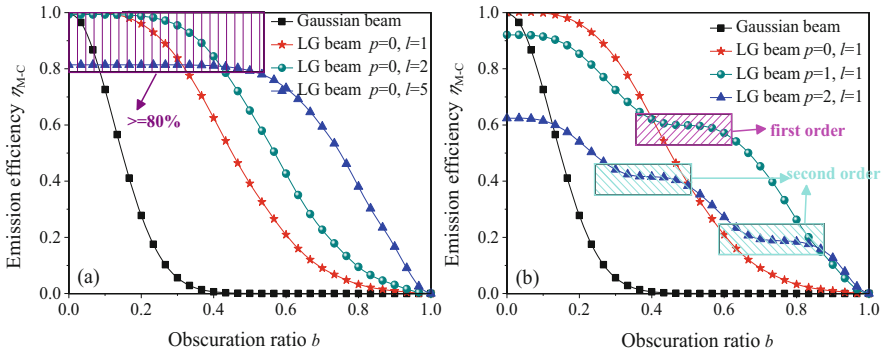


Fig. 48.54 Emission efficiency curve of an LG beam through a Maca-antenna system. (**a**) Different topological charges; (**b**) different radial indices [184]

Research Outlook of OAM Technology for Optical Wireless Communication Systems

With an in-depth study of a structured optical field carrying OAM, the traditional modulation parameters (amplitude, phase, frequency, and polarization state) can no longer meet the demand. The emergence of OAM provides a new type of space-dimensional resource, which opens up new avenues for the development of optical communication.

First, the domestic and foreign development status of OAM-related technologies was outlined, and the advantages and disadvantages were summarized and compared. In addition, the relevant research work of Xi'an University of Technology was introduced, including single-mode and multimode OAM beam generation, partial coherence, vector OAM beam transmission characteristics, crosstalk analysis, wavefront correction, grating aperture detection, separation and detection based on

the coordinate-transformation method in OAM-multiplexed communication systems, and practical applications.

With the advent of global communication integration, the establishment of 6G technology-based land, sea, air, sky, and micro-communication networks and their cross-scene intelligent optical communication networks is a major trend in communication development. Future OAM-related technologies for optical wireless communication systems can progress in the following areas.

1. Generate new OAM beam arrays and study their transmission characteristics to improve the beam energy at the transmitting end, thus enhancing the transmission efficiency of optical wireless communication systems.
2. Develop a geometric coordinate transformation of optical OAM toward compactness and miniaturization and combine it with improved iterative algorithms to improve the resolution and efficiency of the light spot, after separation at the receiving end.
3. Combine MIMO equalization and channel-coding techniques with OAM techniques, which can further mitigate the effect of turbulence on the beam.
4. Establish a multilayer phase diffraction optical neural network to realize a device that integrates the generation, multiplexing, demultiplexing, and detection of OAM beams, then, add iterative algorithms and error functions for optimization and evaluation. Thus, the integrated application of OAM technology in optical wireless communication can be advanced more quickly.

References

1. Allen L, Beijersbergen M W, Spreeuw R J C, et al. Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes [J]. Physics Review A, 1992, 45(11): 8185–8189.
2. Bozinovic N, Yue Y, Ren Y, et al. Terabit-scale orbital angular momentum mode Division multiplexing in fibers [J]. Science, 2013, 340(6140): 1545–1548.
3. Paterson L, MacDonald M P, Arlt J, et al. Controlled rotation of optically trapped microscopic particles [J]. Science, 2001, 292(5518): 912–914.
4. Ruffato G, Massari M, Romanato F. Multiplication and division of the orbital angular momentum of light with diffractive transformation optics [J]. Light: Science & Applications, 2019, 8(1): 113–125.
5. Leach J, Jack B, Romero J, et al. Quantum correlations in optical angle-orbital angular momentum variables [J]. Science, 2010, 329(5992): 662–665.
6. Ke Xizheng. Optical wireless communication [M]. Beijing: Science Press, 2016: 10–13.
7. Ke Xizheng, Wu Jiali, Yang Shangjun. Research progress and prospect of atmospheric turbulence for Optical Wireless Communication [J]. Chinese Journal of Radio Science, 2021, (03): 323–339.
8. Ke Xizheng, Xie Yanchen, Zhang Ying. Orbital Angular Momentum Measurement of Vortex Beam and Its Performance Improvement [J]. Acta Optica Sinica, 2019, 39(01): 258–264.
9. Barnett S M, Allen L. Orbital angular momentum and nonparaxial light beams [J]. Optics Communications, 1994, 110(5–6): 670–678.
10. Coullet, P. Gil, L. Roca F. Optical vortices [J]. Optics Communication, 1989, 73(5): 403–408.

11. Lee A J, Omatsu T, Pask H M. Direct generation of a first-Stokes vortex laser beam from a self-Raman laser [J]. Optics Express, 2013, 21(10): 12401–12409.
12. Lee A J, Zhang C Y, Omatsu T, et al. An intracavity, frequency-doubled self-Raman vortex laser [J]. Optics Express, 2014, 22(5): 5400–5409.
13. Heckenberg, N R, McDuff, R, Smith, C. P. Generation of optical phase singularities by computer-generated holograms [J]. Opt.Lett, 1992, 17(3): 221–223.
14. Li Yanming, Kim Jihwan and Escuti M J. Orbital angular momentum generation and mode transformation with high efficiency using forked polarization gratings [J]. Appl. Opt, 2012, 51(34): 8236–8245.
15. Vijayakumar A, Rosales G C, Rai M R, et al. Generation of structured light by multilevel orbital angular momentum holograms [J]. Opt. Express, 27(5): 6459.
16. Beijersbergen M W, Allen L, H.E.L.O. van der Veen, et al. Astigmatic laser mode converters and transfer of orbital angular momentum [J]. Opt. Commun., 1993, 96(1): 123–139.
17. Shvedov V G, Izdebskaya Y V, Alekseev A N, et al. The formation of optical vortices in the course of light diffraction on a dielectric wedge [J]. Technical Physics Letters, 2002, 28(3): 256–259.
18. Forbes A, Dudley A, McLaren M. Creation and detection of optical modes with spatial light modulators [J]. Adv. Opt. Photonics, 2016, 8(2): 200–227.
19. K. Saitoh, Y. Hasegawa, N. Tanaka, and M. Uchida. Production of electron vortex beams carrying large orbital angular momentum using spiral zone plates [J]. J. Electron. Microsc, 2012, 61(3): 171.
20. Hu X B, Rosales G C. Generation and characterization of complex vector modes with digital micromirror devices: a tutorial [J]. J. Opt, 2022, 24: 034001.
21. Liu R, Li F and Padgett M J et al. Generalized photon sieves:fine control of complex fields with simple pinhole arrays [J]. Optica, 2015, 2(12): 1028–1049.
22. Devlin R C, Ambrosio A and Rubin N A et al. Arbitrary spin-to-otbital angular momentum conversion of light [J]. Science, 2017, 358(6365): 896–901.
23. Dorrah A. H, Rubin N A, et al. Structuring total angular momentum of light along the propagation direction with polarization-controlled meta-optics [J]. Nat. Commun, 2021, 12: 6249.
24. Wu W, Sheng Z and Wu H. Design and application of flat spiral phase plate [J]. Acta Phys.Sin. 2019, 68(5): 054102(1–6).
25. Brunet C, Vaity P, Messaddeq Y, et al. Design, fabrication and validation of an OAM fiber supporting 36 states [J]. Optics Express, 2014, 22(21): 26117–26127.
26. Pidishety S, Pachava S, Gregg P, et al. Orbital angular momentum beam excitation using an all-fiber weakly fused mode selective coupler [J]. Opt. Lett, 2017, 42(21): 4347–4350.
27. Hassan M M, Kabir M A, Hossain M N, et al. Numerical analysis of circular core shaped photonic crystal fiber for orbital angular momentum with efficient transmission [J]. Appl. Phys. B, 2020, 126(9): 145–153.
28. Israk M F, Razzak M A, Ahmed K, et al. Ring-based coil structure photonic crystal fiber for transmission of Orbital Angular Momentum with large bandwidth: Outline, investigation and analysis [J]. Optics Communications, 2020, 473(15): 126003.
29. Cognée K G, Doeleman H M, Lalanne P, et al. Generation of pure OAM beams with a single state of polarization by antenna-decorated microdisk resonators [J]. ACS photonics, 2020, 7(11): 3049–3060.
30. Butt M A, Khonina S N, Kazanskiy N L. Device performance of standard strip, slot and hybrid plasmonic micro-ring resonator: A comparative study [J]. Waves Random Complex Media, 2021(6): 2397–2406.
31. Pyragaite V, Stabinis A. Interference of intersecting singular beams [J]. Optics Communications, 2003, 220(4): 247–25.
32. Galvez E J. Composite optical vortices formed by collinear Laguerre-Gauss beams [J]. Proceedings of SPIE, 2006, 6131:19–26.

33. Baumann S M, Kalb D M, Macmillan L H, et al. Propagation dynamics of optical vortices due to Gouy phase [J]. Optics Express, 2009, 17(12): 9818–9827.
34. Yang Dexing, Wang Hao, He Dongsheng, et al. Complex Vortices Formed by Coxial Superposition of Circular and Planar Vortices [J]. ACTA PHOTONICA SINICA, 2009, 38(7): 1683–1686.
35. Ando T, Matsumoto N, Ohtake Y, et al. Structure of optical singularities in coaxial superpositions of Laguerre-Gaussian modes [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2010, 27(12): 2602–2612.
36. Fang Guijuan, Sun Shunhong, Pu Jixiong. Experimental study of fractional double vortex beams [J]. Acta Physica Sinica, 2012, 61(6): 266–272.
37. Zhu L, Wang J. Simultaneous generation of multiple orbital angular momentum (OAM) modes using a single phase-only element [J]. Optics Express, 2015, 23(20): 26221–26244.
38. Huang S, Miao Z, He C, et al. Composite vortex beams by coaxial superposition of Laguerre–Gaussian beams [J]. Optics & Lasers in Engineering, 2016, 78(1): 132–139.
39. Mateusz S, Jan M, Ireneusz A. Generation of composite vortex beams by independent Spatial Light Modulator pixel addressing [J]. Optics Communications, 2020, 463: 125341(1–8).
40. Kumar Nirjhar, Arora Ankit and Krishnan Ananth. Single-shot generation of composite optical vortex beams using hybrid binary fork gratings [J]. Opt. Express, 2021, 29: 33703–33715.
41. Kotlyar V V, Kovalev A A, Savelyeva A A. Topological charge of a superposition of identical parallel single-ringed Laguerre-Gaussian beams [J]. Computer Optics, 2022, 46(2): 184–188.
42. Wang S, Zhang S L, Qiao H C, et al. Direct generation of vortex beams from a double-end polarized pumped Yb:KYW laser [J]. Optics Express, 2018, 26(21): 26925–26939.
43. Bisson J F, Li J, Ueda K, et al. Generation of Laguerre-Gaussian modes in Nd:YAG laser using diffractive optical pumping [J]. Laser. Phys. Lett, 2005, 2(7): 327.
44. Omatsu, T., Miyamoto, K. & Lee, A. J. Wavelength-versatile optical vortex lasers [J]. J. Opt, 2017, 19(12): 123002.
45. A.Ito, Y.Kozawa,andS.Sato. Generation of hollow scalar and vector beams using a spot-defect mirror [J]. J. Opt. Soc. Am. A, 2010, 27(9): 2072–2077.
46. Kano K, Kozawa Y and Sato S. Generation of purely single transverse mode vortex beam from a He–Ne laser cavity with a spot-defect mirror [J]. Int. J. Opt, 2012, 359141.
47. Okida M, Omatsu T, Itoh M, et al. Direct generation of high power Laguerre-Gaussian output from a diode-pumped Nd:YVO4 1.3-μm bounce laser [J]. Optics Express, 2007, 15(12): 7616–7622.
48. Thomas G M, Minassian A, Sheng X and Damzen M J. Diode-pumped Alexandrite lasers in Q-switched and cavity-dumped Q-switched operation [J]. Opt. Express, 2016, 24(24): 27212–26.
49. Forbes A. Controlling light's helicity at the source: orbital angular momentum states from lasers [J]. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci, 2017, 375(2087): 20150449.
50. Veinhard M, Séverine Bellanger, Daniault L, et al. Orbital Angular Momentum beams generation from 61 channels Coherent Beam Combining femtosecond Digital Laser [J]. Optics Letters, 2021, 46(1): 25–28.
51. Gerchberg R.W, Saxton W D. A practical algorithm for the determination of the phase from image and diffraction plane pictures [J]. Optik, 1972, 35: 237–246.
52. Hsu W.F, Lin S C. Iternative pixelwise approach applied to computer-generated holograms and diffractive optical elements [J]. Appl. Opt, 2018, 57(1): A189–A196.
53. Guo C S, Liu X, Ren X Y, et al. Optimal annular computer-generated holograms for the generation of optical vortices [J]. Journal of the Optical Society of America A Optics Image Science & Vision, 2005, 22(2): 385–390.
54. ITU-R Document 3J/31-E, "On propagation data and prediction methods required for the design of space-to-earth and earth to-space optical communication systems", Radio- communication Study Group Meeting, Budapest, 2001, p7.

55. Paterson C. Atmospheric Turbulence and Orbital Angular Momentum of Single Photons for Optical Communication [J]. Phys. Rev. Lett., 2005, 94(15): 153901-1–153901-4.
56. Anguita J, Neifeld M, Vasic B. Turbulence-induced channel crosstalk in an orbital angular momentum-multiplexed free space optical link [J]. Applied Optics, 2008, 47(13): 2414–2435.
57. Gbur G and Tyson R K. Vortex beam propagation through atmospheric turbulence and topological charge conservation [J]. J. Opt. Soc. Am. A, 2008, 25(1): 225–237.
58. Tyler G, Boyd R. Influence of atmospheric turbulence on the propagation of quantum states of light carrying orbital angular momentum [J]. Optics Letters, 2009, 34(2): 142–144.
59. Wang T, Pu J X, Chen Z Y. Beam-spreading and topological charge of vortex beams propagating in a turbulent atmosphere [J]. Opt. Commun, 2009, 282(7): 1255–1259.
60. Liu X H, Pu J X. Investigation on the scintillation reduction of elliptical vortex beams propagating in atmospheric turbulence [J]. Opt. Express, 2011, 19(27): 26444–26450.
61. Ding Panfeng, Pu Jixiong. Change of the off-center Laguerre-Gaussian vortex beam while propagation [J]. Acta Physica Sinica, 2012, 61(06): 198–203.
62. Lukin V P, Konyaev P A, Sennikov V A. Beam spreading of vortex beams propagating in turbulent atmosphere [J]. Applied Optics, 2012, 51(10): C84-C87.
63. Aksenov V P, Kolosov V V, Pogutsa C E. The influence of the vortex phase on the random wandering of a Laguerre-Gaussian beam propagating in a turbulent atmosphere: a numerical experiment [J]. Journal of Optics, 2013, 15(4): 4007–4013.
64. Banakh V A, Falits A V. Turbulent broadening of Laguerre-Gaussian beam in the atmosphere [J]. Optics & Spectroscopy, 2014, 117(6):942–948.
65. Ge Xiaolu, Wang Benyi, Guo Chengshan. Beam Broadening of Vortex Beams Propagating in Turbulent Atmosphere [J]. Acta Optica Sinica, 2016, 36(3): 16–22.
66. Filimonov G A, Aksenov V P, Kolosov V V, et al. Fluctuations of the orbital angular momentum of vortex laser beam in turbulent atmosphere: dependence on turbulence strength and beam parameters [C]. Proc.SPIE 10035, 22nd International Symposium on Atmospheric & Ocean Optics: Atmospheric Physics, 2016: 100352F(1–6).
67. Porfirev A, Kirilenko M, Khonina S N, et al. Study of propagation of vortex beams in aerosol optical medium [J]. Appl. Opt, 2017, 56(11): E5–E15.
68. Tian Q, Zhu L, Wang Y, et al. The Propagation Properties of a Longitudinal Orbital Angular Momentum Multiplexing System in Atmospheric Turbulence [J]. IEEE Photonics Journal, 2018, 10(1): 1–16.
69. Luo Chuankai, Lu Fang, Miao Zhifang, et al. Propagation and Spreading of Radial Vortex Beam Array in Atmosphere [J]. Acta Optica Sinica, 2019, 39(6): 28–44.
70. Yan Jiawei, Yong Kangle, Tang Shanfa, et al. Comparison of Propagation Characteristics Between Super Gaussian and Gaussian Vortex Beams in Air [J]. Acta Optica Sinica, 2020, 40(2): 17–22.
71. Lazrek M, Hricha Z, Belafhal A. Propagation properties of vortex cosine-hyperbolic-Gaussian beams through oceanic turbulence [J]. Opt Quant Electron, 2022, 54: 172.
72. Wang F, Cai Y, Korotkova O. Partially coherent standard and elegant Laguerre-Gaussian beams of all orders [J]. Optics Express, 2009, 17(25): 22366–22379.
73. Zhang Y, Liu L, Wang F, et al. Average intensity and spectral shifts of a partially coherent standard and elegant Laguerre-Gaussian beam beyond paraxial approximation [J]. Optical and Quantum Electronics, 2014, 46(2): 365–379.
74. Zhang Y, Cui Z, Shi J, et al. Propagation properties of partially coherent Laguerre-Gaussian beams in turbulent atmosphere [J]. Optics & Laser Technology, 2011, 43(4): 741–747.
75. Tang M, Zhao D. Propagation of multi-Gaussian Schell-model vortex beams in isotropic random media [J]. Optics Express, 2015, 23(25): 32766–32776.
76. Wang X, Yao M, Yi X, et al. Spreading and evolution behavior of coherent vortices of multi-Gaussian Schell-model vortex beams propagating through non-Kolmogorov turbulence [J]. Optics & Laser Technology, 2017, 87(25): 99–107.

77. Stahl C S D and Gbur G. Twisted vortex Gaussian Schell-model beams [J]. J. Opt. Soc. Am. A, 2018, 35(11): 1899–1906.
78. Hyde IV M W. Twisted spatiotemporal optical vortex random fields [J]. IEEE Photonics Journal, 2021, 13(2): 1–16.
79. Hricha Z, Halba E M, Lazrek M, et al. Focusing properties and focal shift of partially coherent vortex cosine-hyperbolic-Gaussian beams [J]. Journal of Modern Optics, 2022, 69(14): 779–790.
80. Cheng W, Haus J W, Zhan Q. Propagation of vector vortex beams through a turbulent atmosphere [J]. Optics Express, 2009, 17(20): 17829–17849.
81. Martinez-Herrero R, Prado F. Overall characterization of the polarization structure of radially polarized partially coherent vortex beam [C]. 8th Iberoamerican Optics Meeting and 11th Latin American Meeting on Optics, Lasers, and Application, Porto, Portugal, 2013, 8785:8785AB.
82. Ou J, Jiang Y, He Y. Intensity abd polarization properties of elliptically polarized vortex beams in turbulent atmosphere [J]. Optics & Laser Technology, 2015, 67: 1–7.
83. Cheng K, Zhou Y, Lu G, et al. Vectorial structures of linear-polarized Butterfly-Gauss vortex beams in far zone [J]. Optics Communications, 2018, 415: 107–114.
84. Guo L, Chen Y, Liu X, et al. Vortex phase-induced changes of the statistical properties of a partially coherent radially polarized beam [J]. Optics Express, 2016, 24(13): 13714–13735.
85. Eyyuboğlu H T. Scintillation behaviour of vortex beams in strong turbulence region [J]. Journal of Modern Optics, 63:21, 2374–2381.
86. Tang M, Zhao D, Li X, et al. Propagation of radially polarized multi-cosine Gaussian Schell-model beams in non-Kolmogorov turbulence [J]. Optics Communications, 2018, 407: 392–397.
87. Joshi Stuti, Khan S N, Manisha, et al. Coherence-induced polarization effects in vector vortex beams [J]. Opt. Lett, 2020, 45(17): 4815–4818.
88. Joshi S, Khan S N, Senthilkumaran P, et al. Statistical properties of partially coherent polarization singular vector beams [J]. Physical Review A, 2021, 103(5): 053502.
89. Manisha, Joshi Stuti, Khan S N, et al. Investigation of partially coherent vector vortex beams with non-isotropic states of spatial correlation [J]. Opt. Express, 2022, 30: 32230–32243.
90. Arora Gauri, Joshi Stuti, Singh Hanuman, et al. Perturbation of V-point polarization singular vector beams [J]. Optics & Laser Technology, 2022, 158: 108842.
91. Starikov F A, Kochemasov G G, Koltygin M O, et al. Correction of vortex laser beam in a closed-loop adaptive system with bimorph mirror [J]. Optics Letters, 2009, 34(15): 2264–2266.
92. Garanin S G, Manachinsky A N, Starikov F A, et al. Phase correction of laser radiation with the use of adaptive optical systems at the Russian Federal Nuclea Center—Institute of Experimental Physics [J]. Optoelectronics, Instrumentation and Data Processing, 2012, 48(2): 134–140.
93. Zhao S M, Leach J, Gong L Y, et al. Aberration corrections for free-space optical communications in atmosphere turbulence using orbital angular momentum states [J]. Optics Express, 2012, 20(1):452–461.
94. Jesacher A, Schwaighofer A, Fürhapter S, et al. Wavefront correction of spatial light modulators using an optical vortex image [J]. Optics Express, 2007, 15(9): 5801–5808.
95. Baranek M, Behal J, Bouchal Z. Optimal spiral phase modulation in Gerchberg-Saxton algorithm for wavefront reconstruction and correction [C]//Thirteenth International Conference on Correlation Optics. SPIE, 2018, 10612: 71–78.
96. Xie G, Ren Y, Huang H, et al. Phase correction for a distorted orbital angular momentum beam using a Zernike polynomials-based stochastic-parallel-gradient-descent algorithm [J]. Optics Letters, 2015, 40(7): 197–1200.
97. Xu Zihao. Research on High-resolution Liquid Crystal Adaptive Optics Technique with Phase Diversity [D].Changchun: University of Chinese Academy of Sciences, 2018.
98. Zhai Y W, Fu S Y, Zhang J Q, et al. Turbulence aberration correction for vector vortex beams using deep neural networks on experimental data [J]. Optics Express, 2020, 28(5): 7515–7527.

99. Zhan H C, Wang L, Wang W N, et al. Experimental analysis of adaptive optics correction methods on the beam carrying orbital angular momentum mode through oceanic turbulence [J]. Optik, 2021, 240: 166990(1–8).
100. Jiang Y S, Wang S H, Zhang J H, et al. Spiral spectrum of Laguerre–Gaussian beam propagation in non-Kolmogorov turbulence [J]. Opt. Commun., 2013, 303(16): 38–41.
101. Liu D, Wang Y, Wang G, et al. Intensity properties of flat-topped vortex hollow beams propagating in atmospheric turbulence [J]. Optik – International Journal for Light and Electron Optics, 2016, 127(20): 9386–9393.
102. Xu Y, Tian H, Feng H, et al. Propagation factors of standard and elegant Laguerre-Gaussian beams in non-Kolmogorov turbulence [J]. Optik, 2016, 127(22): 10999–11008.
103. Hricha Z, Lazrek M, Yaalou M, et al. Propagation of vortex cosine-hyperbolic-Gaussian beams in atmospheric turbulence [J]. Optical and Quantum Electronics, 2021, 53(7): 1–15.
104. Li Jinhong, Lv Baida. Comparative study of partially coherent vortex beam propagations through atmospheric turbulence along a uplink path and adownlink path [J]. Acta Physica Sinica, 2011, 60(7): 286–292.
105. Wang D, Wang F, Cai Y, et al. Evolution properties of the complex degree of coherence of a partially coherent Laguerre-Gaussian beams in turbulent atmosphere [J]. Journal of Modern Optics, 2012, 59(4): 372–380.
106. Chen Y, Wang F, Zhao C, et al. Experimental demonstration of a Laguerre-Gaussian correlated Schell-model vortex beam [J]. Optics Express, 2014, 22(5): 5826–5838
107. Zhang Y, Liu L, Zhao C, et al. Multi-Gaussian Schell-model vortex beam [J]. Physics Letters A, 2014, 378(9): 750–754.
108. Zhang L, Deng D, Chen X, et al. Propagation of a radially polarized partially coherent rotating elliptical cosine-Gaussian beam with vortices in anisotropic turbulence [J]. Applied Physics B, 2019, 125(5): 79–90.
109. Hu Juntao, Ma Haixiang, Li Xinzhong, et al. Characteristics of Concentric Vectorial Perfect Vortex Mode [J]. Acta Optica Sinica, 2019, 39(1): 245–251.
110. Ma Shengjie, Hao Shiqi, Zhao Qingsong. Wavefront Distortion Correction of Vortex Beam Based on Improved SPGD Algorithm [J]. Acta Optica Sinica, 2021, 41(06): 7–13.
111. Wu Ran, Chen Jun, Fu Gangkun. Experimental Study of Multi-Mode High-Order Vortex Beam Wander [J]. Acta Optica Sinica, 2022, 42(04): 182–189.
112. Wang Mingjun, Zhang Jialin, Wang Zhuyu, et al. Scattering of Laguerre-Gaussian Vortex Beams by Underwater Suspended Spherical Algal Particle Swarms [J]. Acta Optica Sinica, 2022, 42(18): 199–206.
113. Leach J, Padgett M J, Barnett S M, et al. Measuring the orbital angular momentum of a single photo [J]. Phys. Rev. Lett., 2002, 88(25): 257901(1–4).
114. Vasnetsov M V, Torres J P, Petrov D V, et al. Observation of the orbital angular momentum spectrum of a light beam [J]. Optics Letters, 2003, 28(23): 2285–2287.
115. Ghai D P, Senthilkumaran P, Sirohi R S. Single-slit diffraction of an optical beam with phase singularity [J]. Optics and Lasers in Engineering, 2009, 35(23): 123–126.
116. Silva J G, Jesus-Silva A J, Alencar M A R C, Hickmann J M and Fonseca E J S. Unveiling square and triangular optical lattices: a comparative study [J]. Opt. Lett. 2014, 39(4): 949–952.
117. Emile O and Emile J. Young's double-slit interference pattern from a twisted beam [J]. Optica, 2014, 117(1): 487–491.
118. Zhou H, Shi L, Zhang X, et al. Dynamic interferometry measurement of orbital angular momentum of light [J]. Optics Letters, 2014, 39(20): 6058–61.
119. Fu S, Wang T, Zhang S, et al. Integrating 5×5 Dammann gratings to detect orbital angular momentum states of beams with the range of −24 to +24[J]. Applied Optics, 2016, 55(7): 1514–1517.
120. Li Y, Deng J, Li J, et al. Sensitive Orbital Angular Momentum (OAM) Monitoring by Using Gradually Changing-Period Phase Grating in OAM-Multiplexing Optical Communication Systems [J]. IEEE Photonics Journal, 2016, 8(2): 1–6.

121. Taira Y, Zhang S K. Split in phase singularities of an optical vortex by off-axis diffraction through a simple circular aperture [J]. Optics Letters, 2017, 42(7): 1373–1376.
122. Acevedo C, Moreno Y T, Guzmán Á, et al. Far-field diffraction pattern of an optical light beam with orbital angular momentum through of a rectangular and pentagonal aperture [J]. Optik, 2018, 164: 479–487.
123. Li Y X, Han Y P, Cui Z W, et al. Measuring the topological charge of vortex beams with gradually changing-period spiral spoke grating [J]. IEEE Photonics Technology Letters, 2020, 32(2): 101–104.
124. Kotlyar V V, Kovalev A A and Porfirev A P. Astigmatic transforms of an optical vortex for measurement of its topological charge [J]. Appl. Opt, 2017, 56(14): 4095–4104.
125. Chen J, Chen X, Li T and Zhu S. On-chip detection of orbital angular momentum beam by plasmonic nanogratings [J]. Laser Photon. Rev, 2018, 12(8): 1700331(1–5).
126. Li J, Zhang M, Wang D, Wu S and Zhan Y. Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication [J]. Opt. Express, 2018, 26(8): 10494–10508.
127. Giordani T, Suprano A, Polino E, et al. Machine learning-based classification of vector vortex beams [J]. Physical review letters, 2020, 124(16): 160401.
128. Rasouli S, Fathollazade S, Amiri P. Simple, efficient and reliable characterization of Laguerre-Gaussian beams with non-zero radial indices in diffraction from an amplitude parabolic-line linear grating [J]. Optics Express, 2021, 29(19): 29661–29675.
129. Jiang X, Wang A, Yao J, et al. Measuring high-order multiple vortex beams with fork-shaped grating [J]. Optik, 2022, 257: 168742.
130. Berkhout G C, Lavery M P, Courtial J, et al. Efficient sorting of orbital angular momentum states of light [J]. Physics. Review. Letters., 2010, 105(15): 153601(1–4).
131. Mirhosseini M, Malik M, Shi Z, et al. Efficient separation of the orbital angular momentum eigenstates of light [J]. Nature Communications, 2013, 4(1):1–6.
132. Ruffato G, Massari M, Parisi G, et al. Test of mode-division multiplexing and demultiplexing in free-space with diffractive transformation optics [J]. Optics Express, 2017, 25(7): 7859–7868.
133. Ruffato G, Massari M, Romanato F. Compact sorting of optical vortices by means of diffractive transformation optics [J]. Optics Letters, 2017, 42(3): 551–554.
134. Li C, Zhao S. Efficient separating orbital angular momentum mode with radial varying phase [J]. Photonics Research, 2017, 5(4): 267–270.
135. Wen Y, Chremmos I, Chen Y, et al. Spiral transformation for high-resolution and efficient sorting of optical vortex modes [J]. Physical Review Letters, 2018, 120(19): 193904.
136. Ruffato G, Massari M, Girardi M, et al. Non-paraxial design and fabrication of a compact OAM sorter in the telecom infrared [J].Optics Express, 2019, 27(17): 24123–24146.
137. Fontaine N K, Ryf R, Chen H, et al. Laguerre-Gaussian mode sorter [J]. Nature Communications, 2019, 10(1): 1–7.
138. He L, Lin Z, Wen Y, et al. An inverse design method combining particle swarm optimization and wavefront matching method for multiplane light conversion [C]//Frontiers in Optics, Optica Publishing Group, 2020: FM7D.5.
139. Huang Z, Wang P, Liu J, et al. All-optical signal processing of vortex beams with diffractive deep neural networks [J]. Physical Review Applied, 2021, 15(1): 014037(1–14).
140. Gao C, Dai K, Zhong L, et al. Measuring OAM states of light beams with gradually-changing-period gratings [J]. Optics Letters, 2015, 40(4): 562–565.
141. Fu S, Wang T, Gao Y, et al. Diagnostics of the topological charge of optical vortex by a phase-diffractive element [J]. Chinese Optics Letters, 2016, 14(8): 080501–080505.
142. Zheng S, Wang J. Measuring Orbital Angular Momentum (OAM) States of Vortex Beams with Annular Gratings [J]. Science Reports, 2017, 7(1): 1–9.
143. Ke Xizheng. Generation, transmission, detection and application of vortex beam [M].Beijing: Science Press, 2018: 10–13.

144. Ke Xizheng, Li Yaxing. Experimental Study on Fractional Laguerre Gaussian Beam with Orbital Angular Momentum [J]. Laser & Optoelectronics Progress, 2015, 52(08): 88–94.
145. Ke Xizheng, Ge Tian. Experiment on Generation of Vortex Light with Few-Mode Fiber [J]. Chinese Journal of Lasers, 2017, 44(11): 182–189.
146. Ke Xizheng, Chen Yun, Zhang Ying. Design and Analysis of a Transmission Fiber with High-Order Orbital Angular Momentum Mode [J]. Laser & Optoelectronics Progress, 2019, 56(02): 32–50.
147. Ke X Z, Wang S S. Design of photonic crystal fiber capable of carrying multiple orbital angular momentum modes transmission [J]. Optics and Photonics Journal, 2020, 10(4): 49–63.
148. Ke Xizheng, Xu Junyu. Interference and Detection of Vortex Beams with Orbital Angular Momentum [J]. Chinese Journal of Lasers, 2016, 43(09): 198–203.
149. Ke Xizheng, Xue Pu. Generation of Orbital Angular Momentum superpositions and its test [J]. Infrared and Laser Engineering, 2018, 47(04): 63–68.
150. Ke Xizheng, Shi Xinyu. Experimental study on the superposition of high-order radial Laguerre-Gaussian beams [J]. Infrared and Laser Engineering, 2018, 47(12): 335–341.
151. Ke X Z, Zhao J. Analysis on characteristic of Laguerre-Gaussian beams with topological charges of arithmetic progression [J]. Optik, 2019, 183: 302–310.
152. Ke X Z, Zhang L. Study on orthogonal superposition generation method of double-ring vortex beams [J]. European physical journal D, 2020, 74(3): 1–6.
153. Lv Hong, Ke Xizheng. Scattering of a beam with orbital angular momentum by a single sphere [J]. Acta Physica Sinica, 2009, 58(12): 8302–8308.
154. Lv Hong, Ke Xizheng. Research on the Beam with Orbital Angular Momentum Used in Encoding and Decoding of Optical Communication [J]. Acta Optica Sinica, 2009, 29(02): 331–348.
155. Ke Xizheng, Lu Ning, Yang Qinling. Research of transmission characteristics of single-photon orbital angular momentum [J]. Acta Physica Sinica, 2010, 59(09): 6159–6163.
156. Lv Hong, Ke Xizheng. Effects of aperture on the orbital angular momentum of hollow vortex beam [J]. Science China Physics, Mechanics & Astronomy, 2011, 41(08): 932–941.
157. Li Tie, Chen Juan, Ke Xizheng, et al. Study of orbital angular momentum entangled photons entanglement in atmospheric channel [J]. Acta Physica Sinica, 2012, 61(12): 276–283.
158. Chen Juan, Ke Xizheng, Wang Tiecheng. Joint detection probability of orbital angular momentum entangled photon pairs [J]. Chinese Journal of Quantum Electronics, 2013, 30(04): 430–438.
159. Ke Xizheng, Chen Juan, Yang Yiming. Study on orbital angular momentum of Laguerre-Gaussian beam in a slant-path atmospheric turbulence [J]. Acta Physica Sinica, 2014,63(15): 9–15.
160. Ke Xizheng, Guo Xinlong. Realization of optical phase information encode by using orbital angular momentum of light beam [J]. Chinese Journal of Quantum Electronics, 2015, 32(01): 69–76.
161. Ke Xizheng, Ning Chuan, Wang Jiao. Crosstalk analysis of orbital angular momentum-multiplexed state under atmospheric turbulence [J]. Infrared and Laser Engineering, 2018, 47(11): 402–408.
162. Ke X Z, Wang J, Wang M J et al. Diffraction characteristics of a Laguerre-Gaussian beam through a Maksutov-Cassegrain optical system [J]. Applied Optics, 2018, 57(10): 2570–2576.
163. Wang J, Ke X Z, Wang M J. Evolution of phase singularities of Laguerre–Gaussian Schell-model vortex beam [J]. Journal of Optics, 2020, 22(11): 5602–5607.
164. Ke X Z, Chen S T. Reconfigurable orbital angular momentum exchange assisted by dual-area mirror [J]. Optical Engineering, 2020, 59(11): 116104(1–8).
165. Ke X Z, Li J. Using MCMA-MUK algorithm to suppress crosstalk in orbital angular momentum multiplexing communication system [J]. Optical Review, 2021, 28(4): 331–341.
166. Ke Xizheng, Wang Chaozhen. Intensity Distribution of the Partially Coherent Vortex Beams Propagating in Atmospheric Turbulence [J]. Laser & Optoelectronics Progress, 2016, 53(11): 83–91.

167. Wang J, Ke X Z, Wang M J. Influence of source parameters and atmospheric turbulence on the polarization properties of partially coherent electromagnetic vortex beams [J]. Applied Optics, 2019, 58(24): 6486–6494
168. Ke Xizheng, Wang Xiayao. Experimental Study on the Correction of Wavefront Distortion for Vortex Beam [J]. Acta Optica Sinica, 2018, 38(03): 204–210.
169. Ke X Z, Cui N M. Experimental research on phase diversity method for correcting vortex beam distortion wavefront [J]. Appl. Phys. B, 2020, 126(4): 1–11.
170. Wang J, Wang M J, Ke X Z, et al. Direct Determination of Laguerre-Gauss Vortex Beams' Topological Charges by Aperture Diffraction Pattern [C]. Progress in Electromagnetics Research Symposium(PIERS), Toyama, Japan, 2018: 2155–2159.
171. Ge Tian. Theoretical and Experiemntal Research on the Use of Optical Fiber to Generate Vortex Beam [D]. Xi'an: Xi'an University of Technology, 2018.
172. Chen Yun. Design and Research of Transmission Fiber with High Order Orbital Angular Momentum Mode [D]. Xi'an: Xi'an University of Technology, 2019.
173. Wang Shasha. Research and Design of the Fiber for Generating Multiple Orbital Angular Momentum Modes [D]. Xi'an: Xi'an University of Technology, 2020.
174. Tian W, Zhang H, Zhang X. A circular photonic crystal fiber supporting 26 OAM modes [J]. Optical Fiber Technology, 2016, 30: 184–189.
175. Bai X, Chen H, Ma Y, et al. Circular-lattice photonic crystal fiber with square air holes supporting 58 OAM modes [C]//Progress in Electromagnetics Research Symposium-Spring (PIERS). IEEE, 2017: 3105–3108.
176. Li Yaxing. The Experimental Study of Vortex Beam and Comparative Transmission Characteristic of Orbital Angular Momentum [D]. Xi'an: Xi'an University of Technology, 2016.
177. Xu Junyu. The Experimental Study of Important Technologies of OAM Beam Generation and Control [D]. Xi'an: Xi'an University of Technology, 2017.
178. Xue Pu. Separation and Detection of Orbital Angular Momentum of Vortex Beam [D]. Xi'an: Xi'an University of Technology, 2017.
179. Zhang Lei. Study on The Generation and Transmission Characteristics of Multi-Ring Vortex Beams [D]. Xi'an: Xi'an University of Technology, 2020.
180. Zhao Jie. Distribution Characteristics of Light Field by Superposition of Multiple Laguerre-Gaussian Beams [D]. Xi'an: Xi'an University of Technology, 2019.
181. Shi Xinyu. The Characteristics of The Superposition of Laguerre-Gaussian Beams [D]. Xi'an: Xi'an University of Technology, 2018.
182. Wang Chanzhen. The Intensity Distribution of Partially Coherent Vortex Beam Propagation in The Atmospheric [D]. Xi'an: Xi'an University of Technology, 2017.
183. Ning Chuan. Orbital Angular Momentum Multiplexing and Crosstalk Analysis under Atmospheric [D]. Xi'an: Xi'an University of Technology, 2018.
184. Wang Jiao. Propagation and Diffraction Properties of Partially Coherent Vortex Beams in Atmospheric Turbulence [D]. Xi'an: Xi'an University of Technology, 2020.
185. Wang Tiecheng. Study of Photons Orbital Angular Momentum Entangled in The Quantum Communication [D]. Xi'an: Xi'an University of Technology, 2012.
186. Yang Yiming. Study on The Orbital Angular Momentum of Photons entangled Characteristics [D]. Xi'an: Xi'an University of Technology, 2014.
187. Yang Qinling. The Wireless Laser Communications for Beam Based on Orbital Angular Momentum [D]. Xi'an: Xi'an University of Technology, 2010.
188. Li Jin. Research on Crosstalk Suppression of OAM-MIMO System in Atmospheric Turbulence Channel [D]. Xi'an: Xi'an University of Technology, 2021.
189. Wang Xiayao. Research on Adaptive Optics Correction Technique of Vortex Beams [D]. Xi'an: Xi'an University of Technology, 2018.
190. Cui Namei. Experimental Research on Correcting Wavefront Distortion of Vortex Beam by Phase Diversity Method [D]. Xi'an: Xi'an University of Technology, 2020.

191. Lv Hong. Research on The Orbital Angular Momentum of Vortex Beam for Space Photon Communication [D]. Xi'an: Xi'an University of Technology, 2011.
192. Xie Yanchen. Research on Measurement Technology of Orbital Angular Momentum of Vortex Beam [D]. Xi'an: Xi'an University of Technology, 2019.
193. Guo Xinlong. Study on Orbital Angular Momentum of Beams Encoding and Transmission Characteristics [D]. Xi'an: Xi'an University of Technology, 2015.
194. Chen Shengtao. Research on Vortex Beam Information Exchange And Channel Reconstruction [D]. Xi'an: Xi'an University of Technology, 2021.

Research Progress on Aircraft Relay Optical Wireless Communication

49

Contents

Introduction 1830
Overview of UAV Relay Optical Communication Systems 1830
UAV Relays 1830
Relay Collaboration Technology 1831
Research Status of UAV Relays 1837
Status of Foreign Research 1837
Status of Domestic Research 1841
Key Technologies for UAV Relay Optical Communication 1846
Channel-Influencing Factors and Their Suppression Techniques 1846
High-Precision Capture, Alignment, and Tracking Technology 1851
Modulation and Demodulation Techniques 1853
Lightweight and Miniaturization Design Technology 1856
Non-Co-axial Aircraft-Relay Technology 1857
Application of Optical Wireless Communication to UAV Relaying 1858
Summary 1859
References 1861

Abstract

Aircraft relay can quickly build end-to-end communication networks, enhance the coverage and transmission capability of communication systems, strengthen signal strength and increase data transmission speed, and improve the reliability of communication systems. This chapter summarizes the current research status of UAV relay in optical wireless communication at home and abroad and introduces the basic principles and key technologies of UAV relay in optical wireless communication. Finally, it outlooks the application and development direction of UAV relay in optical wireless communication.

Keywords

UAV · Optical wireless communication · Relay collaboration · Key technologies · Research progress

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_49

Introduction

Optical wireless communication uses light waves to transmit data, typically using visible light, infrared, and other light waves for communication [1]. It transmits digital information by modulating the intensity and frequency of light, among other parameters [2]. It has a high transmission rate and unregulated operating frequency bands [3].

In the optical wireless communication process, the transmission distance and signal quality are affected by factors, such as terrain, buildings, and weather [4]. Therefore, when the signal source and target are far apart or obstacles are present, relaying is required to ensure that the signal can reach the target. Relaying can forward the signal to a greater distance and extend the signal coverage, thus improving the communication efficiency and reliability [5].

With the advancement of unmanned aerial vehicle (UAV) technology and the reduction of its production cost [6], using UAVs as relays for optical communication has become a new communication method [7]. Signals are transmitted from the ground to a high-altitude UAV, and then the UAV relay transmits the signals to the target location, thus enhancing and extending the signal [8–10].

When UAVs communicate via radio frequency (RF), they can interfere with existing ground-based wireless networks, which makes it challenging to implement airborne RF links [11]. Optical wireless communication links can provide higher bandwidth and security than traditional RF links [12], and their application in UAV relaying has been the subject of extensive attention and research. UAV-relaying technology can also enhance the coverage and bandwidth of free-space optical (FSO) communication networks, which is of great importance for optical wireless communication.

This chapter introduces the status of domestic and foreign research on UAV relays in optical wireless communication, describes the principles of the UAV relay system, including relay forwarding strategies, relay system models, protocols used in optical UAV relays, and an evaluation index of system performance. It also focuses on key UAV relay technologies in optical communication, such as channel characteristics, acquisition, pointing, and tracking (APT) technology, modulation and demodulation technology, and lightweight system designs. Finally, the application and development trends of UAV relays in optical wireless communication are foreseen.

Overview of UAV Relay Optical Communication Systems

UAV Relays

Nowadays, the UAV industry is developing rapidly and increasing numbers of UAVs are available. Commonly used relay vehicles [13] are fixed-wing UAVs, rotary-wing UAVs, tethered UAVs, and high-altitude platforms [14]. As shown in Fig. 49.1, fixed-wing UAVs have the advantages of long flight time, fast flight speed, long

distance, and a large load; thus, they are chosen for long-range and high-load relay communication. Multi-rotor UAVs have become the first choice for UAV applications, owing to their ability to hover at any time and their lower production costs.

Tethered UAVs have low environmental requirements for takeoff and landing and long working hours and are flexible and easy to deploy. They are widely used as relays for communication in harsh environments or in the event of natural disasters. High-altitude platforms, also known as quasi-static vehicles, are unmanned aerial systems (UAS) in a broad sense, with ultra-high payloads and power. They are often used as relay nodes between satellites and the ground, as a way to reduce the burden of satellite communications.

The four types of UAVs mentioned above are widely used in relay communications. Compared with other high-altitude vehicles, UAVs have the advantages of flexibility and speed and can provide communication support in areas that cannot be covered by ground communication networks. Moreover, UAVs can adjust their position and altitude at any time to achieve the best relaying effect. Therefore, an effective optical communication link can be established between air–ground and air–air by using UAVs as relays.

Figure 49.2 shows the model of a UAV relay communication system, where the UAV relays between two nodes. The data link between the UAVs is divided into an uplink and a downlink. The uplink sends command information from ground console 1 to the UAV to the telemetry link, and the downlink sends telemetry information from the UAV to the ground console to the telemetry link [15].

Because the UAV flies at a higher altitude, the communication link with the ground control terminal can be approximated as a direct vision communication, which meets the optical wireless communication requirements.

Relay Collaboration Technology

In UAV relay collaboration technology, the signals sent by the source node are processed and forwarded by the relay node and finally delivered to the destination node. The relay node forwarding extends the communication distance. Different relay selection techniques and relay location adjustments can obtain better channel environments, and different relay communication protocols can meet the needs of different application scenarios. The system performance should be evaluated after



Fig. 49.1 Main types of relay aircraft [12]. (**a**) Fixed-wing UAV, (**b**) rotor UAV, (**c**) tethered UAV, and (**d**) high-altitude platform

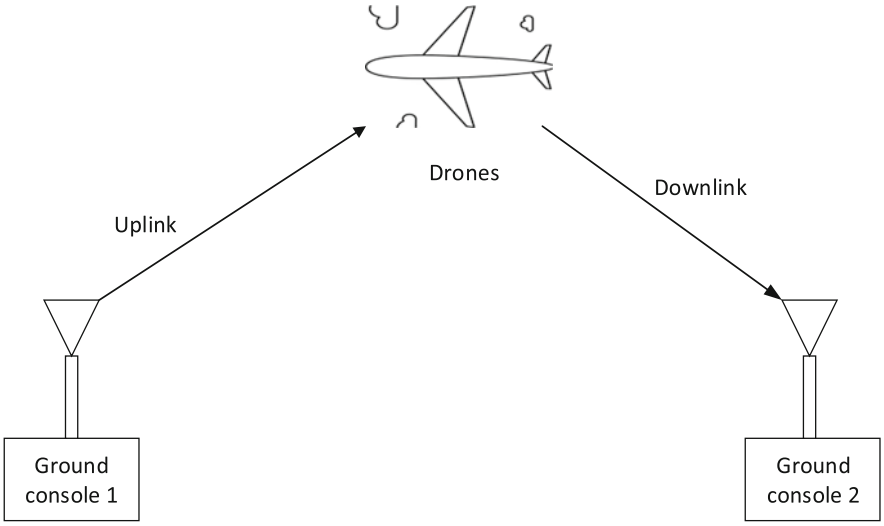


Fig. 49.2 Model of a UAV relay communication system [15]

the UAV completes the relay forwarding. Several relay forwarding strategies, relay system models, communication protocols, and performance evaluation indexes are analyzed below.

Relay Forwarding Strategy

In traditional wireless communication, a relay is used to extend the communication range and increase the line-of-sight transmission. There are three types of relay node forwarding protocols [16]: amplified forwarding, decoded forwarding, and coded forwarding. Figure 49.3 shows a schematic diagram of amplified forwarding, which is a double-hop relay. The double hops are independent of each other and there is no interference. It is better to use amplified forwarding in the case of a high signal-to-noise ratio.

The transmission rates of the UAV relay system under the amplify-and-forward protocol are shown below [17]:

$$R = \log\left(1 + \frac{\mathrm{SNR}_S \mathrm{SNR}_D}{1 + \mathrm{SNR}_S + \mathrm{SNR}_D}\right) \tag{49.1}$$

where SNR_S denotes the signal-to-noise ratio from the source node to the relay node and SNR_D denotes the signal-to-noise ratio from the relay node to the destination node.

Figure 49.4 shows a schematic diagram of decoding and forwarding. The principle is that the signal is demodulated at the relay node and the received signal is checked to determine whether its decoding is correct. If it is correct, it is recoded and forwarded; if not, the relay node will no longer function.

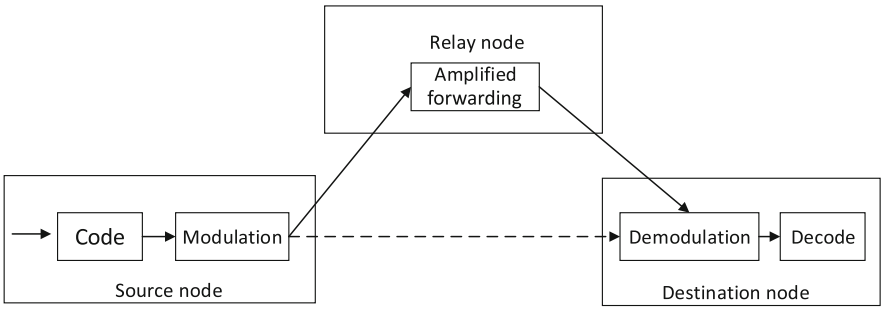


Fig. 49.3 Amplified forwarding schematic [16]

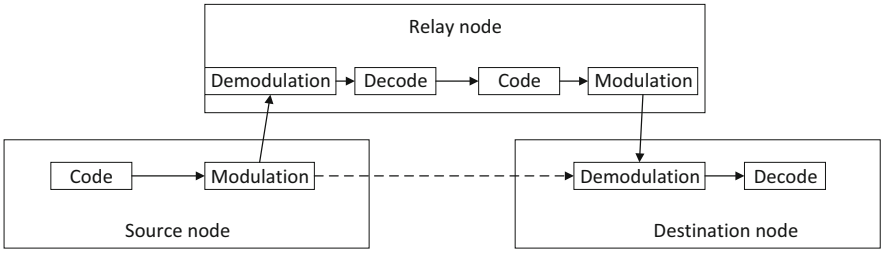


Fig. 49.4 Schematic diagram of decoding and forwarding [16]

The transmission rate of the communication system in a UAV relay system using a decode-and-forward protocol can be described as follows [17]:

$$R = \min\{\log(1 + \mathrm{SNR}_D), \log(1 + \mathrm{SNR}_S)\} \tag{49.2}$$

where SNR_S denotes the signal-to-noise ratio from the source node to the relay node, and SNR_D denotes the signal-to-noise ratio from the relay node to the destination node.

Figure 49.5 shows the schematic diagram of coded forwarding. The principle of this method is similar to decoded forwarding; the difference is that when the decoding result is judged to be wrong, the relay node will encode the incorrect information for forwarding. This method can improve the overall system performance; however, the method is very complicated.

Different relay forwarding strategies have advantages and disadvantages, and they are compared in Table 49.1.

Relay System Models

According to the number of relays forwarded, system models can be divided into single-relay and multi-relay communication, while according to the different relay nodes, relays can be divided into double-hop single-relay and multi-hop multi-relay. Figure 49.6 shows a double-hop single relay. When the communication path from the source node to the destination node is blocked, the relay node can be used for

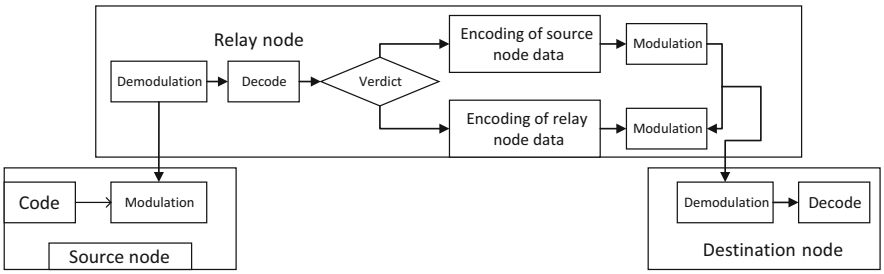
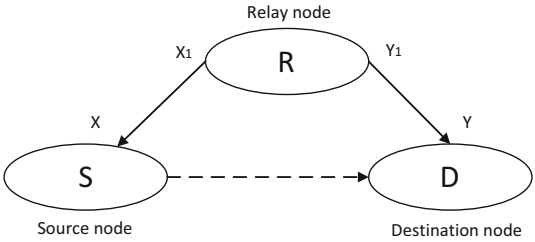


Fig. 49.5 Schematic diagram of coded forwarding [16]

Table 49.1 Comparison of advantages and disadvantages of relay forwarding strategies

Forwarding strategy	Advantages	Disadvantages
Amplify forwarding	More easily improves the signal strength	Unable to handle noise and interference in the signal; the signal quality is easily affected
Decoding and forwarding	Can handle signal noise and interference and perform error correction and error control	The presence of check digits in the checksum makes the transmission less valid and more computationally intensive
Coded forwarding	Encoding and decoding Operations can be performed to optimize signal transmission quality	High complexity and computational effort; less application in practical scenarios

Fig. 49.6 Double-hop single relay [18]



auxiliary communication to expand the communication coverage and improve the performance of the communication system.

S is the source node, D is the destination node, and R is the relay node. The specific process follows. When the S side sends a signal X to the R side, the R side receives signal X_1. The R side executes the prescribed protocol to process certain information on X_1 and sends signal Y_1 to the D side, and the D side receives signal Y. Therefore, the R side plays a crucial role in the communication quality for the entire communication system.

Figure 49.7 shows a multi-hop multi-relay, which adds multiple relay nodes to a two-hop single-relay, further extends the communication coverage, and forwards the

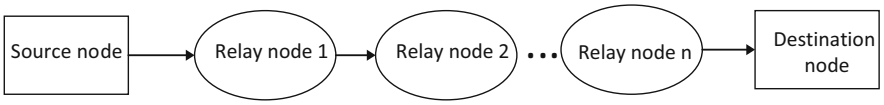


Fig. 49.7 Multi-hop multi-relay [18]

signal multiple times. The disadvantage is that the security rate decreases and the complexity of the entire relay network increases. The communication delay and energy consumption increase exponentially.

Compared with traditional wireless communication networks, relay collaborative communication transmissions have become a trend. Users can choose different relaying methods and forwarding protocols to carry out relay communication, and reasonably selecting a relay node has become the focus of attention. In a UAV relay communication system, the rationality of the selected relay forwarding method and relay node selection are measured by calculating from the interruption probability, signal-to-noise ratio threshold, bit error rate, and channel state information [19].

Relay Communication Protocol

A UAV relay needs to support a variety of communication protocols, including the Transmission Control Protocol/Internet Protocol (TCP/IP) and User Datagram Protocol (UDP) to meet the needs of different application scenarios.

1. TCP/IP protocol: The TCP/IP protocol is a connection-oriented protocol. TCP can provide reliable data transmissions in an IP environment; it can ensure the reliability and integrity of data transmissions and is suitable for scenarios where UAV relays transmit large amounts of data in optical wireless communication.
2. UDP protocol: The UDP protocol is a simple message-oriented transport layer protocol. It is a connectionless protocol with a fast transmission speed but does not guarantee the reliability or integrity of data transmissions. It is suitable for scenarios in which a UAV relay transmits real-time data in optical wireless communication.

Relay Performance Evaluation Metrics

The data transmission capability of UAV relays in optical wireless communication is reflected in the throughput and interruption performance [20], while the reliability and interference immunity of the system are reflected in the average bit error rate, signal-to-noise ratio, and channel capacity.

Interruption Probability

In wireless communication channels, the probability of disruption is generally chosen to reflect the performance of the overall communication system. When the link fades too much, compared to the source, it causes the output instantaneous signal-to-noise ratio of the destination node to be much smaller than the specified signal-to-noise ratio threshold, which in turn causes the signal to be interrupted.

Therefore, the probability that the output instantaneous signal-to-noise ratio of the link is below the threshold is defined as the interruption probability [21]. The expression for the interruption probability can be written as

$$P_{\text{out}} = \Pr(\gamma \leq \gamma_{\text{th}}) \tag{49.3}$$

where γ denotes the instantaneous signal-to-noise ratio of the system, γ_{th} denotes the threshold value, and Pr(•) denotes the probability of an event occurring.

Average Bit Error Rate

The average BER is the ratio of the number of erroneous bits occurring during a transmission to the total number of bits, which reflects the accuracy of the data transmitted by the communication system in a certain period of time. In optical wireless communication systems, the BER is affected by various factors, such as environmental factors, optical parameters, and modulation methods, and is usually measured using experiments or simulations. The lower the average BER, the higher the accuracy of the received data.

Assuming that the system signal is coherently detected and binary phase shift keyed, the BER of the optical wireless communication system is [22]

$$P_e = \int_0^{\infty} \frac{1}{\sqrt{\pi}\sigma^2} \exp\left[-\frac{(i_d + 0.5RIA\zeta)}{\sigma^2}\right] di_d \tag{49.4}$$

where σ^2 denotes the noise variance, i_d denotes the baseband signal strength, R denotes the response coefficient of the photodetector, I denotes half of the peak light intensity, A denotes the carrier amplitude, and ζ denotes the modulation index.

Signal-to-Noise Ratio

The signal-to-noise ratio is the ratio between the signal and the noise power. In an optical wireless communication system, the signal-to-noise ratio directly affects the signal quality and distance. The signal-to-noise ratio of the system can be improved by increasing the transmitting power, optimizing the receiver sensitivity, and reducing the environmental interference. The higher the signal-to-noise ratio, the higher the reliability of the data transmission.

The S/N ratio expression of a UAV relay optical wireless communication system is as follows [22]:

$$\text{SNR} = \frac{(IRA\zeta)^2 P_m}{\sigma^2} \tag{49.5}$$

where σ^2 denotes the noise variance, I denotes half of the peak light intensity, A denotes the carrier amplitude, ζ denotes the modulation index, and P_m denotes the subcarrier signal power.

Channel Capacity

The channel capacity refers to the maximum amount of information transmitted per unit of time and is related to the transmitting power, receiver sensitivity, and transmission distance. In optical wireless communication, it will be affected by fading, noise interference, and other factors. Assuming that the optical signal is affected by atmospheric turbulence factors, the average channel capacity in the optical wireless communication system is

$$C = \int_0^\infty B \log_2(1 + \mathrm{SNR}) P(I) dI \tag{49.6}$$

where B denotes the channel bandwidth, SNR denotes the signal-to-noise ratio, and $P(I)$ denotes the probability density function of the optical signal under atmospheric turbulence.

Relay System Throughput

The throughput refers to the number of packets received per unit time by a target node. This metric is often used to indicate the network's ability to transmit packets. The smaller the throughput value, the worse the transmission performance of the network. In communication systems, the system throughput is defined as the amount of data that is correctly transmitted between the sending node and the destination node within a unit of time [16].

Research Status of UAV Relays

UAV relay optical communication combines the high capacity of optical wireless communication and the high mobility of UAVs, which is of great significance to the further development and improvement of UAV relay optical communication technology.

Status of Foreign Research

The application of optical wireless communication technology in aircraft relaying has attracted widespread attention, and countries worldwide are studying its theory and key technologies. They have also achieved different results.

In 2013, Park et al. [23] investigated an asymmetric amplify-and-forward UAV-relaying approach, using a subcarrier intensity modulation scheme to convert the received RF signal into an optical signal. Under the assumption that the RF and optical wireless channels would experience Rayleigh fading and gamma–gamma fading, respectively, they developed an asymmetric radio frequency/free-space optical (RF/FSO) dual-hop relay system where the average error probability was derived. Its results showed that the average error probability of the asymmetric

RF/FSO UAV relay system was lower than that of the RF/RF relay system under a low signal-to-noise ratio, and better system performance could be obtained.

In 2014, Kaadan et al. [24] investigated the effect of instability of hovering multi-rotor UAVs on the system performance in an optical wireless communication link. It was modeled to measure the power characteristics of fiber optic transceivers at different wavelengths and transmission ranges, enabling for the first time, a short-range optical wireless communication link between two multi-rotor UAVs.

In 2018, Dabiri [25] performed the first stochastic analysis of the channel characteristics of a UAV relay all-optical communication link, assuming that the channel underwent gamma–gamma fading. Statistical models for ground-to-UAV, UAV-to-UAV, and UAV-to-ground links were derived, and the outage probability of each link was derived.

In 2018, Alzenad et al. [26] analyzed the drone characteristics of optical wireless communication under different weather conditions. A point-to-point optical wireless communication link was established between the core network and a terrestrial access point for transmitting data traffic using drone relays to improve the system stability. The research results showed that the backhaul link based on optical wireless communication had a higher transmission rate when the base stations were more densely clustered.

In 2018, Fawaz et al. [27] proposed to apply UAV relaying to optical wireless communication by installing an optical wireless communication transceiver and its optical device on a UAV and using this UAV as a removable relay node with caching capability. It was shown that the effects of atmospheric turbulence and fading losses on the optical wireless communication system could be reduced by setting multiple UAV-assisted relays in harsh environments with longer transmission distances.

In 2018, Dong et al. [28] applied the edge computing method to an unmanned aircraft relay optical wireless communication network to support a wireless network architecture to complement the ground infrastructure. Optical wireless communication receivers were carried on both the forward and return links for a rapid response to unexpected events and flexible deployment.

In 2019, Lee [29] investigated a two-hop FSO/RF heterogeneous system based on UAV relaying and used a time-discrete and continuous-convex approximation method to obtain a suboptimal solution for the UAV trajectory, with the maximum throughput rate as the optimization objective. Based on this, a UAV trajectory optimization method based on dichotomous and sequential programming methods was proposed to maximize the flight time.

In 2019, Lee et al. [30] studied the trajectory optimization of fixed-wing UAV relaying in optical wireless communication and used the dichotomous and sequential planning methods to maximize the service time of a vertical wireless backhaul network, based on optical wireless communication.

In 2019, Wu et al. [31] studied UAV-relay-assisted mobile access networks under optical wireless communication. They designed a user quality of service (QoS)-aware algorithm to maximize the number of served mobile users by using UAV-base station and mobile-user association strategies to jointly optimize UAV base-station deployment, mobile-user association, and bandwidth allocation.

In 2019, Safi et al. [32] studied the effect of pointing errors on the link, owing to UAV fluctuations in a UAV relay optical wireless communication system. Because the fluctuations of hovering UAVs caused beam drift and angle-of-arrival undulations, the beam was tracked using quadrant detectors to find the optimal detector size that minimized the tracking error. Closed expressions for the tracking error and BER were derived, and Monte Carlo simulations were performed to verify the accuracy of the expressions.

In 2020, Usman et al. [33] proposed a hybrid backhaul for mobile base station-assisted cellular networks and derived expressions for the BER of RF, free-space optical, and hybrid backhaul communication. In the hybrid connection, optical wireless communication was selected when the UAV relay was in a fixed position, and RF communication was selected when the UAV relay was flying from one position to another. It was also demonstrated that this hybrid backhaul network outperformed the single optical wireless communication and single RF communication scenarios.

In 2020, Ansari [34] used a hovering UAV as a relay for optical wireless communication between a ground base station and a user. To increase the hovering time of the UAV, an optical wireless communication receiver was placed on the UAV, along with a solar panel. The optical wireless receiver received a portion of the beam from the ground base station and demodulated the data carried by the beam. The solar panel received the remaining beam and absorbed the energy carried by the beam. The results of the study showed that the method effectively increased the hover time and data transmission rate.

In 2021, Lee [35] investigated the throughput maximization of FSO/RF-hybrid UAV-assisted mobile relaying and designed effective trajectories for UAV relaying under different weather conditions, considering buffering and delay. The throughput maximization problem under buffer constraints was divided into delay-limited and delay-tolerant transmissions, and its effective trajectory was obtained by convex quadratic constraint planning to find the optimal buffer size and delay time requirements. It resulted in a significant increase in throughput compared with the conventional static relaying scheme.

In 2021, Dabiri et al. [36] studied the source-node-to-relay and relay-to-destination-node channel models for UAV-assisted optical wireless communication relay links and derived the optimal three-dimensional coordinates and optimal beam pattern for the UAV relay to minimize the disruption probability. Moreover, physical parameters that affected the optimal UAV relay positioning were investigated, such as the UAV directional fluctuation intensity, obstacle height, and location.

In 2021, Safi [37] investigated a multi-rotor hovering UAV in a ground-space free-space optical link. The UAV used a four-quadrant photodetector array. This method was able to increase the field of view at the receiving end, thus effectively reducing the effects of fluctuations. A variable focal length lens was used to adjust the beam size and reduce the impact of angle-of-arrival jitter and aiming error on the system performance.

In 2021, Wang et al. [38] used a UAV relay-assisted optical wireless communication system to mitigate the limiting effect of turbulence-induced atmospheric

scintillation. A serial optical wireless communication decode-and-forward system based on a hovering UAV was investigated to reveal the effects of different factors, such as atmospheric turbulence, pointing errors, and angle-of-arrival undulations, on the system. The beamwidth, field of view, and UAV position were also optimized, and the results showed that the proposed method could significantly improve the system performance.

In 2021, Dabiri et al. [39] investigated the application of UAV relaying to optical wireless communication amplify-and-forward systems. They derived the statistical probability of the end-to-end signal-to-noise ratio under the influence of the directional and positional fluctuations of the UAV relay, position of the UAV relative to the source and target nodes, beamwidth, and turbulence intensity of the source-to-repeater and relay-to-target links, as well as the BER and interruption probability expressions. The results showed that the optimal UAV height for the minimum interruption probability was determined by the beamwidth at the target, and the optimal value of the source-to-relay node link length increased when the beamwidth at the destination decreased.

In 2022, Wang et al. [40] analyzed the performance of a UAV-assisted dual-hop optical wireless communication relay system, considering the effects of the pointing error, fading loss, and atmospheric turbulence on the source-to-relay and relay-to-destination links. The source-to-relay and relay-to-destination links were assumed to satisfy the Malaga and gamma–gamma distributions, respectively, from which expressions for the probability density function, cumulative distribution function, end-to-end outage probability of the relay system, and average BER were derived. The results showed that in the weak turbulence case, the system performance mainly depended on the turbulent environment, while in the strong turbulence case, the system performance was more sensitive to the position offset.

In 2022, Janji et al. [41] used a UAV relay optical communication system for small cell (SC) backhaul links, in which UAVs with optical wireless communication devices were used as backhaul relay nodes. They proposed to use the visible map algorithm and shortest-path method to find the location of the UAV relay station to ensure direct-view communication between consecutive hops. The final results showed that this method could effectively improve the system gain.

In 2022, Qu et al. [42] studied an integrated air–air, air–ground network in an UAV-assisted free-space optical communication system. Assuming that the RF communication and optical wireless communication followed the shadowing and Malaga fading distributions, respectively, the outage probability, average BER, and traversal capacity for different modulation methods were derived. The simulation results indicated that the system performance was severely degraded when the turbulence conditions or pointing error became large. Moreover, the system with outlier detection performed better than the system with intensity modulation.

In 2023, Lee [43] proposed an FSO/RF satellite–UAV integrated hybrid for nonterrestrial networks using deep reinforcement learning methods to maximize the total end-to-end throughput, while minimizing the UAV energy consumption. Their simulation results indicated that the method could significantly increase the

throughput compared to fixed terrestrial relaying; however, the energy consumption increased.

In 2023, Singh et al. [44] investigated the performance of an optical wireless communication system during decode-and-forward for a hovering UAV relay, considering the effects of path loss, atmospheric turbulence, aiming error, and angle-of-arrival undulation. Using direct detection techniques, a closed-form expression for the point-to-point channel gain was derived, and the interruption performance of a serial multi-hop system was analyzed on this basis.

The development of UAV relaying in optical wireless communication has evolved from individual UAVs to UAV clusters, from fixed-wing UAVs to rotary-wing UAVs, from air–ground links to air–air links, and from FSO/FSO links to FSO/RF links, and the research has evolved from theoretical derivations to simulations to practical applications. The system performance of UAV relay optical communication systems is studied by considering different relay forwarding methods and different fading losses, which significantly improve the system throughput, energy consumption, and other aspects (Table 49.2).

Status of Domestic Research

In the 1990s, China began to study optical wireless communication and at the beginning of the twenty-first century began to develop UAV relay technology. Only in recent years UAV relay and optical wireless communication technology have been combined; however, they were rapidly fully developed and researched, and research results were achieved.

In 2017, An et al. [45] studied the communication condition of a UAV air–ground optical wireless communication system in different weather, in terms of the maximum received power, maximum power variation range, Ethernet packet loss rate, and stability of transmitting high-definition video at a fixed distance and time. Field experiments showed that when the weather temperature rose, the turbulence affected the communication-power variation range. In rainy weather, the power variation range was smaller, and the smaller the power variation, the lower the Ethernet packet-loss rate and the better the communication quality. It provided experience for air–ground long-distance information transmissions by UAV relays in optical wireless communication systems.

In 2018, Xu et al. [46] designed a UAV airborne laser communication device because of the demand for high-precision equipment for UAV-relayed optical wireless communication. It had small aperture reception, adaptive adjustment of the light-range divergence angle, coaxial-signal optical transceiver, and omnidirectional automatic scanning capture tracking. The results showed that the device had the advantages of a small size, low power consumption, high scanning and capturing tracking efficiency, shortened communication link establishment time, and stable data transmissions. This provided a high-precision device for the development of optical wireless communication in UAV relays.

Table 49.2 Foreign research results

Year	Researcher	Research results
2013	Park	Asymmetric amplify-and-forward RF/FSO UAV relay system outperformed an RF/RF UAV relay.
2014	Kaadan	Short-range optical wireless communication between two multi-rotor UAVs achieved for the first time
2018	Dabiri	First stochastic analysis of channel characteristics of all-optical communication links for UAVs
2018	Dong	Applied edge computing to optical wireless communication networks for UAV relays
2019	Lee	Study of a two-hop FSO/RF heterogeneous system to obtain a suboptimal solution for UAV trajectories
2019	Lee	Maximized flight time obtained by trajectory optimization of fixed-wing UAVs
2019	Wu	Designed user quality-of-service perception algorithm to maximize the number of mobile users of the service
2019	Safi	Found the optimal detector size to minimize tracking errors, based on the impact of the pointing error on a ground-to-UAV optical wireless communication link
2020	Usman	By comparing the BER of a single RF link, a single optical wireless communication link, and a hybrid link, it was concluded that the performance of the hybrid link was better.
2020	Ansari	Effectively increased the hovering time of a hovering drone by placing solar panels and optical wireless communication receivers on it
2021	Dabiri	Modeled the channel from the source node to the relay node and the relay node to the destination node to derive the optimal three-dimensional coordinates
2021	Safi	Used a four-quadrant photodetector array to increase the field of view at the receiver and variable focal-length lens to adjust the beam size
2021	Wang	Optimized beamwidth and field of view by analyzing the UAV relay optical communication system under decode-and-forward, considering the effects of atmospheric turbulence, pointing error, and arrival angle undulation on the system
2021	Dabiri	Developed expressions for the interruption probability and BER of a UAV relay optical communication system under amplification-and-forwarding, where the beam width at the target could determine the optimal UAV altitude
2022	Wang	By analyzing the UAV relay optical communication system under decode-and-forward, the outage probability and average BER were derived by considering the effects of fading loss, atmospheric turbulence, and pointing error on the system
2022	Janji	Used UAV relays equipped with optical wireless communication devices for cellular networks to determine the location of UAVs using visible map algorithms and the shortest path method to ensure direct vision communication
2023	Lee	A deep learning-based UAV relay–optimization algorithm was proposed to maximize the throughput when using UAVs as relays between satellites and the ground

In 2018, Xu et al. [47] studied UAV relay optical wireless communication under an amplify-and-forward protocol to maximize the end-to-end throughput of a UAV relay system by power allocation, as well as UAV trajectory. An iterative algorithm was proposed for optimizing the power allocation and UAV trajectory at the source and relay nodes. The results of the study showed that the method could significantly improve the system throughput.

In 2019, Yuan [48] studied the channel characteristics of a two-hop optical communication system with a UAV relay and a multi-hop hybrid RF/optical communication system and derived the probability density functions of the alignment errors from the ground end to UAV and UAV to UAV during the communication process. Because of the maneuverability of UAVs, their coverage and path loss were considered, from which the optimal flight altitude of UAVs with a minimum disruption probability and the effects of different alignment error coefficients on the system performance were derived.

In 2019, Yang et al. [49] studied a multi-hop hybrid RF/FSO communication system, where two UAVs were selected as relays. The outage probability of the system and the optimal UAV altitude were derived by considering the pointing error and path loss. The results showed that the UAV flight altitude had a significant impact on the system performance.

In 2019, Gu [50] studied link load balancing and minimum network delays for UAV-relayed free-space optical communication. The network topology was modeled considering power, link reliability, and link capacity, so that the network capacity could be maximized with a minimum number of UAVs. By jointly optimizing the location and resource allocation of UAV relays, the network latency was effectively reduced.

In 2020, Lei et al. [51] investigated a nonorthogonal multiple access communication system based on UAV hybrid RF/FSO, where the data to be transmitted were converted into an optical signal at the transmitter side after modulation and sent to the UAV. The received data were decoded at the UAV relay and the signal was sent to the user, thereby deriving a closed expression for the interruption probability. The results showed that the UAV flight altitude and alignment error, as well as atmospheric turbulence, affected the system performance.

In 2020, Hu [52] studied the influence of external environmental constraints on the performance of UAV relay optical wireless communication and analyzed the influence of the atmospheric channel and transmitter side on the beam transmission. They detected the UAV transmission link length, stability, channel quality, and other indicators, as well as adjusting in real time the transmitting optical power and beam dispersion angle by designing each UAV module. The results showed that the method could improve the link security and reduce the bit-error rate of the UAV downlink and reduce the optical detector saturation and data transmission interruptions.

In 2020, Ma et al. [53] studied a free-space optical serial multi-hop decode-and-forward relay system based on hovering UAVs. They derived the interruption probability expressions for each link from the ground to UAV, UAV to UAV, and UAV to ground, considering the effects of atmospheric attenuation, atmospheric

loss, pointing error, and arrival angle undulation, focusing on optimizing the effect of arrival angle undulation on the field of view of the UAV receiver. The results showed that the derived expressions could accurately assess the system interruption performance and significantly improve the field-of-view performance.

In 2021, Lu et al. [54] investigated the secrecy performance of UAV-assisted FSO communication systems by modeling the channel, considering the effects of path loss, atmospheric turbulence, pointing error, and reach-angle fluctuations. They derived theoretical expressions for the secrecy interruption probability that could be easily handled.

In 2021, Liang [55] studied automatic beam tracking in UAV optical communications. The performance states of different servo motors in the APT system were compared, and a four-quadrant detector offset calculation method under different energy distribution models and the centering algorithm of the through-hole four-quadrant detector were investigated to capture, align, and track the beam signal using the principle of spot detection. A lightweight and rapidly deployable APT system was designed, using a stepper motor, reflector, and four-quadrant detector instruments, which solved the problem of the limited load of a UAV relay for optical wireless communication systems.

In 2022, Xu [21] studied an optical wireless communication system using an unmanned aircraft as a relay. Atmospheric turbulence, aiming error, angle-of-arrival undulation, and receiver-side noise were analyzed, including the performance of the optical wireless communication system under different noise and turbulence intensities; moreover, expressions for the average bit error rate under atmospheric turbulence and noise were given. The interruption performance of the two-hop relay system was then investigated, and performance metrics for the interruption system were derived. The signal-to-noise ratio of the ground-to-UAV link was derived by considering the fading loss when studying the effect of the link between the ground and the UAV.

In 2022, Xu et al. [56] studied UAV relay communication based on a hybrid optoelectronic architecture, using free-space optical technology between the base station and the UAV relay and RF communication technology between the UAV and the user. This meant that the user information collected by the RF link was forwarded to the base station through an optical wireless communication link. An expression for the interruption probability at a high signal-to-noise ratio was derived, considering the effects of the pointing error and atmospheric turbulence. It was concluded that the beamforming scheme could improve the system diversity and enhance the system performance.

In 2022, Ma [57] studied an optical wireless communication system based on a UAV relay, considering the joint effects of atmospheric turbulence, pointing error, angle-of-arrival jitter, etc. On this basis, the interruption performance of a serial multi-hop relay system was analyzed, and the UAV hovering position, link-reception beamwidth, and field-of-view angle were optimized in turn. UAV relay heterogeneous energy efficiency optimization was also studied, and an iterative algorithm was used to obtain the UAV operation trajectory. The results showed that this type of

algorithm could avoid relay-cache overflow and optimize the energy efficiency of the system.

In 2022, Long [58] studied the physical layer security system for a UAV-relayed optical wireless communication system in amplify-and-forward mode. The wireless channel was modeled under direct and nondirect vision conditions, and the outage security probability and security capacity were derived, considering the distance between the UAV relay and the ground, as well as the path loss. Moreover, a new energy harvesting technique was designed to recollect the direct current component that should be filtered out during optical wireless communication, which improved the energy utilization of the system and effectively enhanced the range of the UAV.

In 2022, Ma [59] studied the use of UAVs as relays between satellites and the ground for resource scheduling and constructed a communication model between UAVs and satellites, based on optical wireless communication under the proportional fairness scheduling algorithm. It selected users with better channel conditions for data transmission at each resource scheduling, which significantly improved the throughput of the system. The communication model between the UAV and users based on RF communication was constructed in decode-and-forward mode, and the signal-to-noise ratios of RF communication and optical wireless communication were derived. The research results showed that the method could significantly improve the fairness of resource allocation for users.

In 2022, Han et al. [60] studied air–ground UAV oblique-range laser communication under the influence of atmospheric turbulence, using a multiple-input single-output technique in the uplink and a single-input multiple-output technique in the downlink. They derived BER expressions for the uplink and downlink and investigated the influence of the zenith angle and pointing error on the system. The results showed that the multiple-input single-output and single-input multiple-output techniques both reduced the power and improved the effect of atmospheric turbulence on the system as well as the communication quality. The light-intensity flicker index varied with the zenith angle and should be controlled within 30 degrees. When the ratio of the pointing error to receiver aperture varied, the BER also varied with the beam dispersion angle; consequently, there exists an optimal beam dispersion angle to optimize the BER performance.

In 2023, Li et al. [61] studied the dynamic switching of optical wireless communication and RF communication based on rate matching. They used a satellite as the source node, a high-altitude platform as a relay, and a UAV as the destination node, using optical wireless communication between the satellite and high-altitude platform. The Radar and Algorithm Display System (RADS) was used for dynamic selection between the high-altitude platform and the UAV, which gave preference to optical wireless communication; when the link's instantaneous signal-to-noise ratio was lower than the threshold value, the RF communication was switched. The research results showed that the algorithm reduced the interruption probability and improved the transmission rate.

Domestic development on designing and optimizing UAV relay optical communication systems has been rapid, mainly from optimizing UAV navigation paths,

improving signal processing algorithms and data transmission protocols, and investigating the hardware equipment needed for UAV relays, such as communication modules, control systems, etc. Then, we considered the factors that affected the channel model when UAV relay optical communication was transmitted in the atmosphere, such as atmospheric turbulence, atmospheric attenuation, pointing error, etc., to optimize the system performance (Table 49.3).

The comprehensive domestic and foreign research status indicated that research on optical wireless communication in UAV relays mainly focused on the following aspects:

1. Research on channel influence factors, such as atmospheric turbulence, fading loss, etc.
2. Research on the rapid establishment of data links, such as rapid beam alignment, etc.
3. Optimization of communication equipment, such as structure design, optical path design, etc.
4. Research on transmission technology, such as relay forwarding methods, modulation and demodulation types, etc.

Research on optical wireless communication in UAV relay systems has made some progress in theory and application; however, there are limitations and shortcomings, and they should be studied in depth.

Key Technologies for UAV Relay Optical Communication

UAV relay optical wireless communication is a relay transmission technology based on UAVs carrying optical wireless communication equipment, such as lasers, to communicate with ground targets. It has the advantages of a high transmission rate, long distance, and anti-interference. To build a stable and efficient system, the following key technologies are focused on.

Channel-Influencing Factors and Their Suppression Techniques

Channel Model

Both UAV optical communication systems and RF systems cannot be kept completely open in arbitrary weather conditions or complex environments. The FSO/RF channel has the advantages of both optical and RF communication and can improve link utilization. Therefore, FSO/RF hybrid transmission technology is still a current research hotspot, and the FSO/RF channel is introduced here.

Because of the high transmission rate of optical wireless communication, it is used between the source node and the relay UAV. RF communication is used

Table 49.3 Domestic research results

Year	Researcher	Research results
2017	An	Designed an optical wireless communication system suitable for small UAVs to achieve high-speed long-distance data transmissions to users
2018	Xu	Designed UAV-airborne laser communication equipment, with the advantages of small size, low power consumption, high scanning and capturing tracking efficiency, short communication-link establishment time, stable data transmissions, etc.
2019	Yuan	Used the channel characteristics of double-hop and multi-hop UAV relay optical communications to obtain the optimal flight altitude for UAVs with a minimum outage probability
2019	Yang	Selected two UAVs as relays to obtain the optimal height of the UAV
2019	Gu	Modeled the network topology to obtain the maximum network capacity with the minimum number of drones
2020	Hu	Adjusted the UAV modules, as well as the emitted optical power and emitted beam dispersion angle, to improve the link security and reduce the optical detector saturation
2020	Ma	Optimized the effect of angle-of-arrival undulation on the UAV receiving field-of-view angle, significantly improving the field-of-view performance
2021	Liang	A lightweight and rapidly deployable APT system was designed to solve the problem of limited UAV loads and an automatic tracking algorithm was designed.
2022	Lu	An expression was derived for the secrecy outage probability that could accurately evaluate the physical layer secrecy performance of the system.
2022	Xu	Considering the effect of noise at the receiver side, an expression for the average BER was derived for different noises and turbulence intensities.
2022	Xu	A new beamforming scheme was derived to improve the diversity of the system by studying the UAV relay communication with a hybrid optoelectronic architecture.
2022	Ma	Considering the optimization of UAV relay heterogeneous energy efficiency, an iterative algorithm was used to optimize the system energy.
2022	Long	New energy-harvesting technology was proposed to improve the energy utilization of the system by collecting the DC component that should be filtered out in optical wireless communication through an energy-harvesting device.
2022	Ma	UAVs were used as relays between satellites and the ground for resource scheduling, optical wireless communication was used between UAVs and satellites, and RF communication was used between UAVs and users to improve the fairness of user-resource allocation.
2023	Tang	Multi-input single-output and single-input multi-output technologies for uplink and downlink, respectively, improved the quality of optical wireless communication.
2023	Li	A high-altitude platform was used as a relay between the satellite and the ground, and the RADS algorithm was used to dynamically select the communication link.

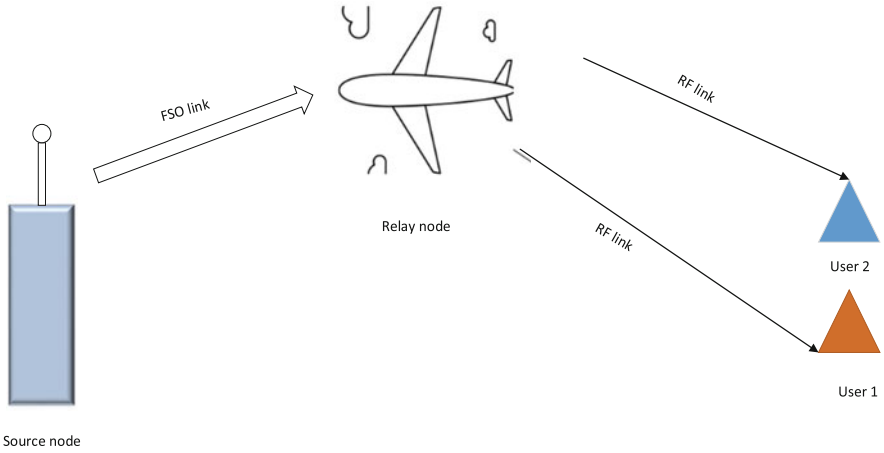


Fig. 49.8 Model of an RF/FSO communication system based on a UAV relay [51]

between the relay UAV and the destination node because of the uncertainty of the user's location and building occlusion [51]. A model of an FSO/RF system based on a UAV relay is shown in Fig. 49.8; it consists of source and relay nodes and users 1 and 2. Optical wireless communication is used between the source node and the relay UAV, and RF communication is used between the relay UAV and the user.

Optical Wireless Communication Link

Suppose users 1 and 2 have a small packet transmission communication service. When the information is sent from the source to the relay drone, the electrical signal can be expressed as $(\alpha_1 s_1 + \alpha_2 s_2)$. s_1 and s_2 denote the packets delivered to the user, α_1 and α_2 are their respective distribution coefficients, and $\alpha_1 + \alpha_2 = 1$. Then, after modulation, the signal is sent to the drone relay, where the direct current signal is removed. After demodulation, the relay can be expressed as

$$y_1 = I\sqrt{p_1}\xi\eta(\alpha_1 s_1 + \alpha_2 s_2) + n_1 \tag{49.7}$$

where $I = h_a h_p$ denotes the optical wireless communication channel coefficient [21], h_a denotes the atmospheric turbulence coefficient, and h_p denotes the pointing error. p_1 denotes the transmit power at the source node, n_1 denotes Gaussian white noise, ξ denotes the conversion efficiency from optical to electrical signals, and η denotes the photoelectric conversion coefficient, whose value is between 0 and 1.

The signal-to-noise ratio of user 1 is

$$\gamma_1 = \frac{p_1(I\xi\eta\alpha_1)^2}{p_1(I\xi\eta\alpha_2)^2 + N_0} \tag{49.8}$$

The signal-to-noise ratio of user 2 is

$$\gamma_2 = \frac{p_1 (I\xi\eta\alpha_2)^2}{N_0} \tag{49.9}$$

RF Communication Link

As can be seen from Eq. 49.7, the signal sent by the UAV relay is y_1. This signal will be sent out as an RF signal by decoding and forwarding; and then the signals received by the users are

$$\gamma_1 = \sqrt{\tau_1}\sqrt{p_2}h_1(\alpha_1 s_1 + \alpha_2 s_2) + n_1 \tag{49.10}$$

$$\gamma_2 = \sqrt{\tau_2}\sqrt{p_2}h_2(\alpha_1 s_1 + \alpha_2 s_2) + n_2 \tag{49.11}$$

where h_1 denotes the channel gain of the RF link between user 1 and the relay, h_2 denotes the channel gain of the RF link between user 2 and the relay, and $h_1 h_2$ follows the Rice distribution. p_2 denotes the transmit power of the relay UAV, n_1 and n_2 denote the Gaussian white noise in the user channel, and τ denotes the path loss attenuation factor.

The signal-to-noise ratios at the user side are the following:

$$\gamma_3 = \frac{\tau_1 p_2 (h_1\alpha_2)^2}{\tau_2 p_2 (h_1\alpha_2)^2 + N_0} \tag{49.12}$$

$$\gamma_4 = \frac{\tau_2 p_2 (h_2\alpha_2)^2}{\tau_2 p_2 (h_2\alpha_1)^2 + N_0} \tag{49.13}$$

where N_0 denotes the power spectral density, γ_3 denotes the signal-to-noise ratio of user 1, and γ_4 denotes the signal-to-noise ratio of user 2.

Channel Characteristics

When a drone is used as a relay for optical communication, the free-space optical signal in the system is affected by atmospheric refraction, absorption, and scattering, degrading the system performance [18]. Several factors affecting the communication system and their suppression methods are described below.

Path Loss

The path loss refers to the decrease in signal strength during signal propagation. It is mainly due to the effects of scattering, reflection, and refraction when the signal is propagating in air, resulting in a gradual weakening of the signal energy. Under the free-space model, the path loss of the signal is [62]

$$L_P = 20\log\left(\frac{4\pi d}{\lambda_c}\right)^2 = 20\log\left(\frac{4\pi d f_c}{c}\right)^2 \tag{49.14}$$

where d denotes the distance between the source and destination nodes, λ_c denotes the wavelength of the transmitted signal, f_c denotes the signal frequency, and c denotes the speed of light. From the above equation, we can see that the energy loss of the signal is proportional to the signal frequency and transmission distance; an excessively high signal frequency or excessively far transmission distance will lead to excessive signal energy loss. UAVs fly at lower altitudes and have smaller energy loss and lower time delays than high-orbiting satellite communication systems.

Doppler Effect

The Doppler effect is caused by the motion of the transmitting and receiving ends. Because there is relative motion between the two when the UAV and ground terminal communicate, it causes the phase and frequency of the received signal to be out of phase with the expected path. The Doppler effect can be expressed as [62]

$$f = \frac{v_c}{\lambda}\cos\beta(t) = f_c\frac{v(t)}{c}\beta(t) \tag{49.15}$$

where f denotes the Doppler shift, $v(t)$ denotes the velocity of the moving object, $\beta(t)$ denotes the angle between the UAV and the direction of ground motion, and f_c is the frequency of the transmitted signal.

Multipath Effect

Owing to the complex wireless propagation environment, the signal in the transmission process often follows more than one transmission path, so the receiver receives the signal at different times. The phases of these signals will also superimpose each other to form interference, resulting in signal distortion. Multipath propagation also causes time delay expansion, so that the signal pulse width of the received signal will appear to broaden, resulting in intersymbol interference.

Atmospheric Turbulence

Atmospheric turbulence may cause the laser pulse signal to spread in time and space and reduce the received optical power. This leads to a reduction of the average signal-to-noise ratio of the communication system and an increase of the communication bit error rate and the number of transient communication interruptions, which have a great impact on the stability and reliability of the communication [63].

Aiming Error

The aiming error is mainly caused by phenomena, such as the photodetector offset and building shaking [62], which mainly include the position offset error and angular tilt error. The aiming error causes the axis between the optical signal transmission and receiving ends to deviate, which reduces the received optical power and limits

the performance of the optical wireless communication system. Therefore, the effect of the aiming error should be considered in a long-range optical wireless signal transmission with UAVs as relays.

When a Gaussian beam is emitted from the transmitter, the aiming error coefficient for the receiver incidence angle is [64]

$$h_{p|\theta_d} = \int_{\pi r^2} I(r - r_1, z)\cos(\theta_d)dr \tag{49.16}$$

where $I(r, z)$ denotes the normalized light intensity distribution, z denotes the transmission distance, and θ_d denotes the incidence angle of the UAV receiver.

Arrival Angle Undulation

The angle-of-arrival undulation is due to the nonstop change in the orientation of the UAV while hovering, the instability of the signal power received by the target node, the jitter in the focusing plane, etc., which leads to a serious decrease in the reliability of the optical wireless communication system.

Assuming that the angle of incidence at the UAV receiver with a smaller viewing angle is θ_d, the θ_d approximation follows the Rayleigh distribution with a probability density function of [65]

$$f_{\theta_d}(\theta_d) = \frac{\theta_d}{\sigma_0^2}\exp\left(-\frac{\theta_d}{2\sigma_0^2}\right) \tag{49.17}$$

where σ_0^2 denotes the link variance.

Table 49.4 summarizes the causes of the abovementioned UAV relay optical wireless communication channel impact factors, their effects on the system, and the suppression techniques.

High-Precision Capture, Alignment, and Tracking Technology

Through analyzing the UAV relay channel model, it is clear that to build a stable and reliable system for effective information transmission between the ground and UAV, the accuracy of the capture, alignment, and tracking technology must be high. Moreover, owing to their small size and lightweight, UAVs are more influenced by airflow, which will affect the flight attitude; hence, fast and accurate beam alignment is the core of UAV airborne laser communication.

To effectively transmit large amounts of information and efficiently reconstruct the wireless laser communication system for stable acquisition, pointing, and tracking (APT), the APT system is critical. APT is divided into three main stages [55].

1. Capture. To connect the UAV to the optical communication link at the ground end, when the optical receiver has not yet captured the beacon light, the loop is

Table 49.4 Channel-influencing factors and their suppression techniques

Influencing factors	Causes	Causes' impact	Suppression technology
Path loss	The signal is scattered, reflected, and refracted when propagating in air.	Gradual weakening of signal energy	Increase transmitting power and optimize antenna design
Doppler effect	Relative motion exists when the UAV and ground terminal are communicating.	Affects the phase and frequency of the received signal	Signal preprocessing, adaptive filtering, or optical path compensation
Multipath effect	The signal has multiple transmission paths in the transmission process, causing the receiving end to receive the signal at different times. The phases superimposed on each other form interference.	Causes signal distortion, time-delay expansion, and pulse spreading	Diversity reception or channel equalization
Atmospheric turbulence	Random changes in the atmospheric temperature and humidity cause changes in the refractive index.	The pulse spreads, received optical power decreases, average signal-to-noise ratio decreases, and communication BER increases	Adaptive optic technology compensates for signal wavefront distortion. Increase optical antenna aperture to suppress optical signal flicker.
Aiming error	This is caused by phenomena, such as photodetector deflection and building shaking.	Reduced received optical power	Build more accurate target models, higher-precision optical sensors, and attitude control systems.
Arrival-angle undulation	This is caused by constant changes in the drone orientation while hovering, instability of the signal power received by the target node, geometric link loss, and jitter in the focus plane.	Reduced reliability of optical wireless communication system	Adaptive modulation technology and adaptive optical path compensation

opened to scan the entire uncertain area information. The scanning stops once the signal light has been detected in the field of view, to complete the capture and lay the foundation for subsequent alignment and tracking [19].

2. Alignment. The UAV optical receiver is aligned with the laser transmitter at the ground end to ensure that the transmitting, as well as tracking related, view axes have high-precision coaxiality. In the process of realizing long-range laser communication, the transmitting end must be overaligned when high-speed relative motion occurs between the transmitting and receiving sides [3].

3. Tracking. The optical signals are tracked and adjusted in real time in the laser-communication link between the UAV and the ground after the alignment is completed. Real-time tracking technology can circumvent the vibration residuals of the platform and the dynamic lag error caused by the relative motion between the two communicating parties [19].

The UAV relay laser communication terminal is divided into the APT system, laser transmitter, and laser receiver, in terms of system function. The working principle of the APT system [67] is shown in Fig. 49.9. The coarse tracking servo ring is composed of a charge-coupled device (CCD), coarse tracking servo controller, and coarse tracking turntable, which completes the initial pointing of the view axis, open-loop capture, and dynamic coarse tracking.

The larger the tracking field of view of the coarse-tracking ring, the smaller the capture time will be; however, with the increase of the tracking field of view, the spatial resolution of the CCD decreases, affecting the final tracking accuracy and bandwidth of the tracking system. Thus, the tracking and field-of-view accuracy must be optimized.

The fine-tracking servo ring is composed of two-dimensional piezoelectric transducer (PZT) mirrors, high bandwidth spot measurement unit, servo control unit, etc. It has the advantages of a small field of view, high servo bandwidth, and high tracking accuracy and can effectively control the dynamic hysteresis bias and platform residuals, so that the accuracy of the APT system meets the standard communication requirements.

To meet the high accuracy laser link alignment and fast automatic establishment of the communication link in UAV relay optical communication, bidirectional data transmissions are realized at the ground and UAV ends [67].

First, the inertial navigation system is used to accurately measure the current information of the carrier to achieve automatic initial pointing. Then, the scanning and capturing technology is used to guide the signal into the tracking phase with the fiducial light and position detector to drive the gyro-servo turntable to guide the laser into the detector-position field of view. Then, the laser-communication link is established. The gyro-servo turntable can reduce the high-frequency vibration of the UAV carrier due to flight, enabling the system to automatically track in the field of view. Figure 49.10 shows the workflow of the APT system in a UAV relay laser communication transmission.

Modulation and Demodulation Techniques

The intensity modulation method is commonly used in UAV relay optical wireless communication. Intensity modulation transmits information by changing the output power of the laser light source. Intensity modulation can be divided into baseband signal intensity modulation and subcarrier intensity modulation.

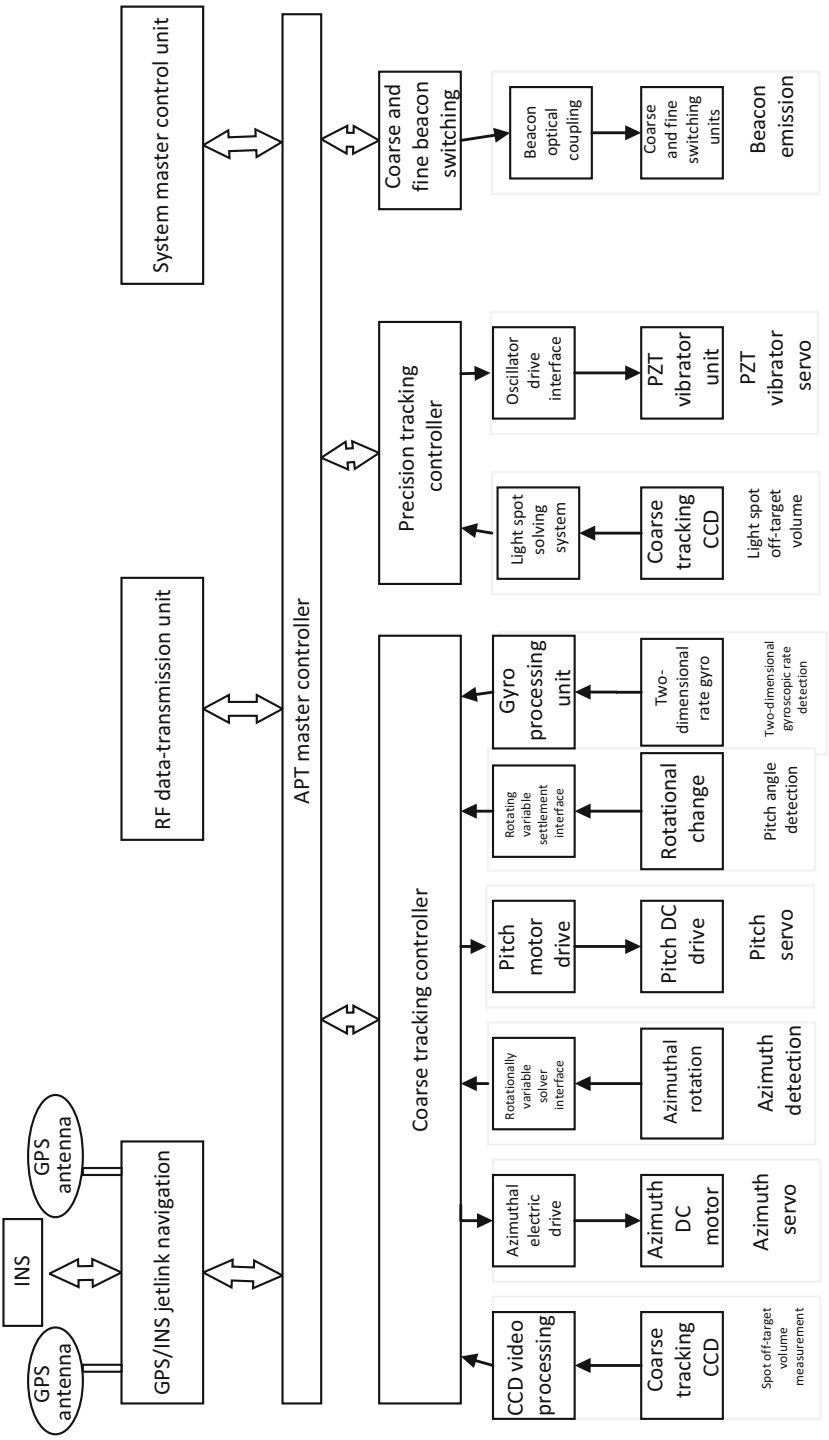


Fig. 49.9 Diagram of the composition of the APT system for laser communication [66]

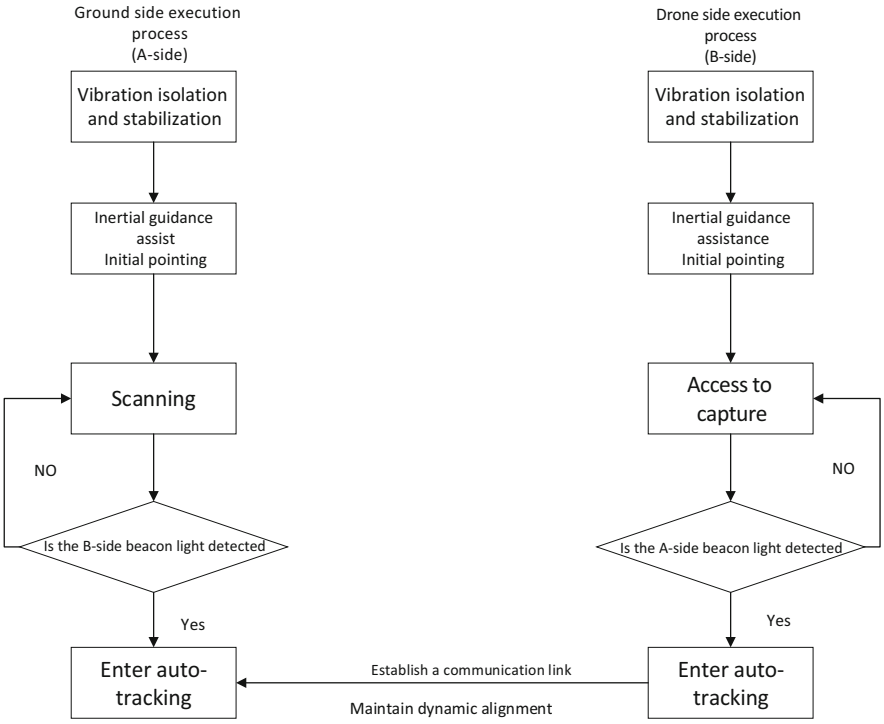


Fig. 49.10 APT system workflow [67]

The most commonly used method in baseband signal intensity modulation is switch keying, which sends light pulses of a given duration when the information bit is 1 and does not send light pulses when the information bit is 0 [48]. The subcarrier intensity modulation technique is also widely used in UAV relay optical wireless communication. It mainly includes phase-shift keying, differential phase-shift keying, and orthogonal phase-shift keying [21].

In terms of signal demodulation, the system needs to demodulate the received signal to extract the valid information. The two demodulation methods commonly used in optical communication are coherent detection and direct detection.

Figure 49.11 shows the block diagram of a coherent detection system. The photodetector accepts the received signal, together with the coherent carrier wave generated by the local oscillator, amplifies and processes it, and finally converts it to the required information signal. Coherent detection has high sensitivity and good noise immunity [47]. Direct detection, on the other hand, directly detects the envelope and converts the intensity-modulated optical signals to electrical signals, which are then amplified [21]. The direct detection technique has the advantages of simplicity and low cost.

Different modulation and demodulation methods have their own advantages and disadvantages, as shown in Table 49.5.

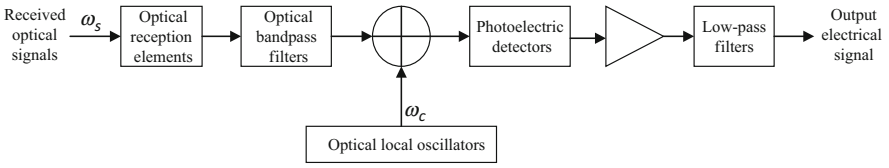


Fig. 49.11 Block diagram of the coherent detection communication system [47]

Table 49.5 Comparison of advantages and disadvantages of modulation and demodulation methods

Method		Advantages	Disadvantages
Modulation	Switch keying	High bandwidth utilization and simple implementation	Poor noise immunity and limited modulation-frequency range
	Subcarrier intensity modulation	Modulated signals have a wide frequency range, allowing transmissions at higher data rates	Requires complex modulation and demodulation circuitry
Demodulation	Correlation testing	Strong anti-interference ability, can achieve higher-precision demodulation	Requires the use of complex circuits and algorithms, with high requirements for signal phase and frequency
	Direct detection	Simple structure, easy to integrate, can increase the transmission rate to complete high-capacity transmissions	Poor anti-jamming capability, low band utilization, and reception sensitivity

Lightweight and Miniaturization Design Technology

While ensuring the speed and stability of the UAV, minimizing its load volume and weight is a key technology in the field of UAV relay optical wireless communication. The smaller the load weight of the UAV, the more flexible and convenient it will be when flying, enabling it to carry more weight. Lightweight miniaturized UAVs can be designed from the following aspects:

1. Use integrated features: By optimizing the selection and layout of components, such as lasers, mirrors, and optical fibers, we can reduce the size of the entire optical-transmission system and integrate the communication equipment with other sensors and navigation systems, thus reducing the number and weight of devices [68].
2. Adopt low-power electronic components: We select low-power electronic components, including power amplifiers, optical receivers, and control circuits, to reduce the energy consumption and required power capacity of the equipment.
3. Adopt efficient optical technology: By adopting efficient optical technology, such as adaptive optics and multi-beam technology, the optical communication

performance can be improved, while reducing the weight and power consumption of the optical communication equipment.

The abovementioned aspects can be used to create lightweight and miniaturized equipment and improve the adaptability and comprehensive performance of UAV relays.

Non-Co-axial Aircraft-Relay Technology

Non-co-axial communication refers to a optical wireless communication process in which the transmitter and receiver cannot be directly aligned. Hence, the antennas of the signal transmitter and receiver must be adjusted so that the transmitter and receiver can be aligned to complete the communication process.

Figure 49.12 shows a schematic diagram of non-co-axial communication [70], in which two-dimensional reflector alignment is used. In the coarse-alignment calibration system from the transmitting antenna to the two-dimensional reflectors, a calibration camera is used to position the calibration, and the transmitting antenna is adjusted according to the feedback image. In a two-dimensional reflectors-to-receiver antenna short-axis fine-alignment system, the two-dimensional reflectors are adjusted according to the feedback from detector 2 for non-co-axial communication.

Non-co-axial laser communication can be well implemented using reflectors, as shown in Fig. 49.13, for the application of non-co-axial communication in UAV relays. A two-dimensional motor turntable and other optical-receiving devices are mounted on the UAV to capture, align, and track the laser signal from the ground

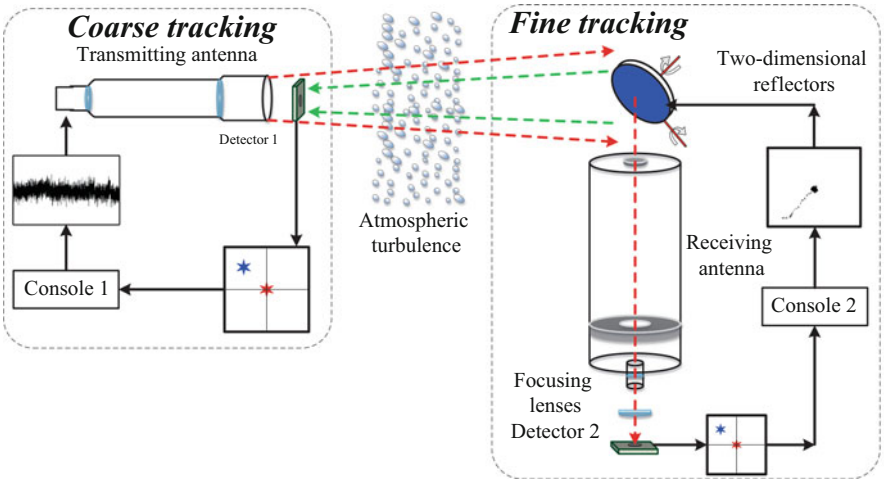


Fig. 49.12 Principle of non-co-axial communication [70]

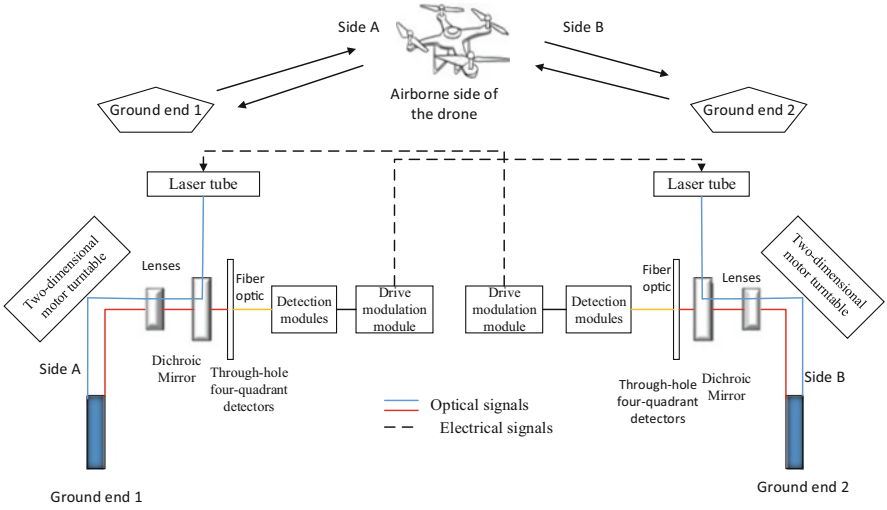


Fig. 49.13 Schematic diagram of a transceiver-integrated UAV relay APT system [54]

end, allowing two ground terminals to form a communication link through the UAV relay communication system [54].

The transmitting end (ground end 1) emits a laser beam signal to the A end of the UAV. It is reflected by the reflector in the two-dimensional motor turntable carried by the relay UAV and then coupled through the lens and the four-quadrant through-hole detector. Then, it enters the optical fiber and is reemitted through the detection module, drive modulation module, and laser tube. Among them, optical fiber is a fiber made of glass and plastic, which acts as a conduction tool for light. Finally, the laser signal is transmitted back to ground end 2 through the dichroic mirror, lens, and two-dimensional motor turntable with the reflector, that is, from the B end to ground end 2. When ground end 1 is the transmitting end, ground end 2 is the receiving end and vice versa.

This method solves the problem of time-consuming beam alignment in conventional optical wireless communication. The method facilitates the establishment of communication links in non-co-axial free space and can help rapidly align laser beams in optical wireless communication using aircraft relays.

Application of Optical Wireless Communication to UAV Relaying

1. Intelligent transportation applications [71]: The UAV relay is used in telematic systems. In recent years, vehicular networking has developed rapidly, which enabled vehicle-to-vehicle, vehicle-to-pedestrian, vehicle-to-infrastructure, and vehicle-to-base station communications. While traffic congestion and obstacles can interfere with the communication links between vehicles and reduce the throughput, using UAVs as relays to assist communication can effectively reduce the problems of network communication, content caching, and security privacy in telematic systems.

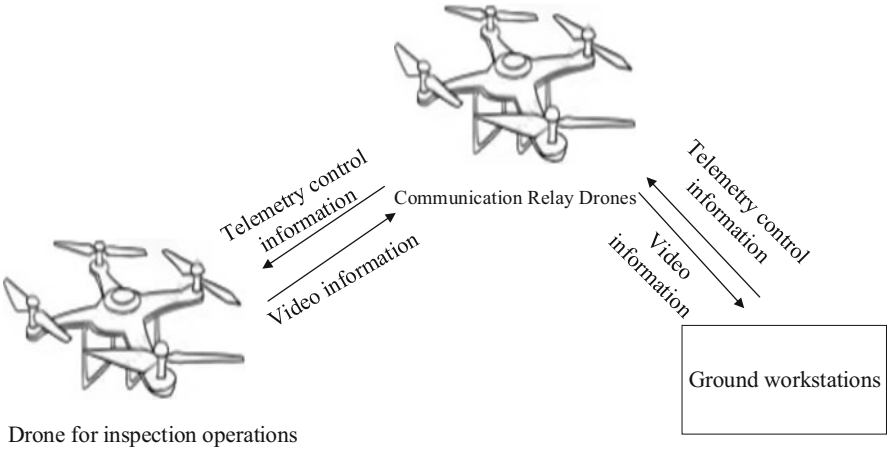


Fig. 49.14 Principle diagram of UAV relay power inspection [73]

2. Environmental-monitoring applications [72]: In response to the forest fires in Australia, domestic scholars built a fire monitoring model based on the interplay of a system safety assessment (SSA) UAV, relay UAV, and emergency operations center. Among them, the SSA UAV cruises to monitor the fire, the relay UAV transmits the data, and the emergency operations center directs firefighters to extinguish the fire. The artificial fish swarm algorithm (AFSA) is used to make specific deployments to the specific location of the emergency operations center, which can minimize economic costs.
3. Electrical power applications [73]: The UAV relay is applied to electric power inspection, and a relay communication system is used to transmit stable signals over long distances. Two small multi-rotor UAVs are selected, one of which carries an infrared HD camera, data transmission, image sending, and other equipment, and the other carries communication-relay equipment.

The overall system setup should consider the aerodynamic shape of the vehicle setup, the communication signal, and the setup of the inspection system. Figure 49.14 shows the schematic diagram of a UAV relay power inspection. The UAV relay achieves over-the-horizon wide-range flight, which expands the communication system coverage and promotes the development of UAV relays in the power industry.

Summary

UAV relays can not only provide long-distance communication but also act as collectors for sending and receiving information. Optical wireless communication has a high transmission rate and good confidentiality and can solve the problem of RF communication congestion. Therefore, combining RF and optical wireless

communication can take advantage of the high mobility of UAVs and improve the stability of existing mobile communication networks.

This chapter introduced the optical wireless communication and UAV relay systems. The research status of UAV relays in optical wireless communication was presented, and the relay-related technologies of UAV relay optical communication systems were introduced. In addition, a brief overview was given of the key technologies of UAV relays in optical wireless communication, such as channel characteristic compensation, APT, modulation and demodulation, lightweight and miniaturized design, etc. Finally, the applications of UAV relays in optical wireless communication are introduced. The future development trends will be mainly in the following areas:

1. UAV relay communication systems will develop in the RF/FSO heterogeneous direction: RF communication is susceptible to interference and cannot communicate efficiently but has a strong ability to penetrate clouds and fog; while free-space optical communication has a strong anti-interference ability, it is easily affected by the atmosphere. Therefore, to maximize the advantages of the different spectrum communication systems, constructing RF/FSO heterogeneous systems for UAV relay communication provides efficient data connections and improves the spectrum utilization and data transmission rate to achieve more efficient, safe, and reliable communication services [74].
2. UAV relay optical communication systems will develop in the direction of artificial intelligence and automation: Nowadays, machine learning and artificial intelligence are developing rapidly. When machine learning is applied to UAV relay optical communication systems [74], it can help to quickly establish communication links, improve the data transmission rate and anti-interference capability, and also enable UAVs to independently select the best relay path. Hence, combining UAV relay optical communication with machine learning and artificial intelligence is a hot spot for future development.
3. The energy consumption of UAV relay optical communication will develop in the direction of energy-saving and green technology: UAV relays can collect energy through solar panels, fuel cells, and other energy collection technologies to ensure their long operation. For example, lithium-polymer batteries are light in weight and have low voltage when the battery is discharged and no memory effect. Lithium-iron-phosphate battery packs are lighter in weight and safer to use and have a greater operating temperature range, allowing ultrasmall and micro-UAVs to have better flight times and payload quality than lithium-polymer batteries.

   Modern composite materials, using glass and carbon fibers and reinforced epoxy resins, have good potential for extending flight time, reducing the weight of UAVs, and increasing their service life. Moreover, UAV relays can be used to ensure the normal equipment operation by distributing the collected energy to optical wireless communication devices and other sensor devices.
4. The relay synergy mechanism of UAV relay optical communication will develop in a more efficient direction: An efficient synergy mechanism should be established between a UAV relay and ground optical communication equipment

to ensure the communication quality and service reliability. The UAV relay optical wireless communication can work in concert with other communication systems, such as satellite and mobile communication, to form a complete communication system. To further improve communication reliability and coverage, future research will enhance the exploration and study of such synergistic mechanisms to improve the performance and efficiency of the entire system

In summary, UAV relay optical wireless communication is an important trend in the future development of communication technology and has broad application prospects in the Internet of Things, intelligent manufacturing, and smart cities.

References

1. Ke Xizheng, Yin Zhiyun. Coding theory in wireless laser communication systems [M]. Beijing: Science Press, 2009:8–10.
2. Ke Xizheng, Deng Lijun. Wireless Laser Communication [M]. Beijing: Science Press, 2016: 1–169.
3. Li Zheng, Xu Dongsheng, Liang Jingyuan, et al. Research on path loss model of outdoor visible light communication [J]. Journal of Xi'an University of Technology, 2023, 23(01):1–10.
4. Liang Jingyuan, Xu Yaxin, Han Meimiao, et al. Research progress of indoor visible light communication heterogeneous fusion system [J]. Optical communication technology, 2023, 47 (01):46–50.
5. Li Zheng, Liao Zhiwen, Liang Jingyuan, et al. Research and Prospect of Atmospheric Turbulence Model [J]. Optical Communication Technology, 2023, 47(03):9–17.
6. Ke Xizheng, Liang Jingyuan, Xu Dongsheng, et al. Research progress of pulse position modulation technology in Optical Wireless Communication [J]. Optoelectronic engineering, 2022, 49 (03):3–21.
7. Ke Xizheng, Qin Huanhuan, Yang Shangjun, et al. Nighttime background light noise model of vehicular visible light communication system [J]. Journal of Radio Wave Science, 2021, 36(06): 986–990.
8. Wang Mingjun, Wei Yafei, Ke Xizheng. Research on laser transmission characteristics between airborne communication terminal and UAV target under complex atmospheric background [J]. According to Chinese Journal of Physics, 2019, 68(09):140–148.
9. Fan Y, Cheng J, Tsiftsis T A. Free-Space Optical Communication with Nonzero Boresight Pointing Errors [J]. IEEE Transactions on Communications, 2014, 62(02):713–725.
10. Yan L S, Wang F, Wu W, et al. Current status and key technologies of unmanned aerial vehicle laser communication payloads [J]. Laser Optoelectron Prog, 2016, 53(08):85–87.
11. Lee H, Sang H L, Quek T, et al. Deep Learning Framework for Wireless Systems: Applications to Optical Wireless Communications [J]. IEEE Communications Magazine, 2019, 57(03): 35–41.
12. Lang Weimin, Tian Shangbao, Li Yuge, et al. Research on the airborne relay communication system [J]. Telecom Express, 2022, 10(03):1–5.
13. Ding Jupeng, Yi Zhiling, Wang Jingtao, et al. Research progress on airborne optical wireless communication for unmanned aerial vehicles [J]. Progress in laser and optoelectronics, 2020, 57(23):40–51.
14. Hou Tianhao, Xing Hongyan., Liu Yang. Status and prospects of multi-rotor UAVs in weather detection [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(10):7–15.
15. Tong Jinghui, Ding Pei, Hua Min, et al. Trajectory optimization of UAV relay network nodes based on reinforcement learning [J]. Communication technology, 2023, 56(01):49–55.

16. He Dan. Research on cooperative relay forwarding strategy optimization and coded cooperation [D]. Zhejiang Hangzhou: Hangzhou Dianzi University, 2017.
17. Fang Jin. Research on trajectory optimization and resource allocation of UAV relay system [D]. Hohhot, Inner Mongolia: Inner Mongolia University, 2022.
18. Zhao Feng. Research on multi-UAV cooperation and location planning for wireless relay [D]. Nanjing, Jiangsu: Nanjing University of Posts and Telecommunications, 2022.
19. Hong Lufeng, Yang Xiaofei, Liu Zhanjun. Cooperative partner selectio5n strategy based on outage probability [J]. Guangdong Communication Technology, 2010, 30 (02):39–42.
20. Hou Jialin. Effectiveness evaluation and analysis of UAV relay transmission system [D]. Xi'an, Shaanxi: Xi'an University of Electronic Science and Technology, 2021.
21. Farid A A, Hranilovic S. Outage capacity optimization for free-space optical links with pointing errors [J]. Journal of Light wave technology, 2007, 25(7):1702–1710.
22. Xu Zixuan. Research on the performance of UAV-assisted Optical Wireless Communication relay system [D]. Shanghai: East China Normal University, 2022.
23. Park J, Lee E, Park G, et al. Performance Analysis of Asymmetric RF/FSO Dual-Hop Relaying Systems for UAV Applications [A]. IEEE Military Communications Conference [C]. San Diego: IEEE. 2013:1651–1656.
24. Kaadan A, Refai H, LoPresti P G. Multielement FSO transceivers alignment for inter-UAV communications [J]. Journal of Lightwave Technology, 2014, 32(24): 4785–4795.
25. Dabiri M T, Sadough S M S, Khalighi M A. Channel Modeling and Parameter Optimization for Hovering UAV-Based Free-Space Optical Links [J]. IEEE Journal on Selected Areas in Communications, 2018, 36(9):2104–2113.
26. Dabiri M T, Khankalantary S, Piran M J, et al. UAV-assisted free space optical communication system with amplify-and-forward relaying [J]. IEEE Transactions on Vehicular Technology, 2021, 70(9):8926–8949.
27. Fawaz W, Abou-Rjeily C, Assi C. UAV-aided cooperation for FSO communication systems [J]. IEEE Communications Magazine, 2018, 56(1):70–75.
28. Dong Y, Hassan M Z, Cheng J, et al. An edge computing empowered radio access network with UAV-mounted FSO fronthaul and backhaul: Key challenges and approaches [J]. Wireless Communications, 2018, 25(3):154–160.
29. J.-H. Lee, K.-H. Park, M.-S, et al. On the Throughput of Mixed FSO/RF UAV-Enabled Mobile Relaying Systems with a Buffer Constraint [A]. IEEE International Conference on Communications [C]. Shanghai: IEEE. 2019:1–6.
30. Lee J H, Park K H, Alouini M S, et al. Free space optical communication on UAV-assisted backhaul networks: Optimization for service time [A]. IEEE Globecom Workshops [C]. Waikoloa: IEEE. 2019:1–6.
31. Lee J H, Park K H, Ko Y C, et al. Throughput maximization of mixed FSO/RF UAV-aided mobile relaying with a buffer [A]. IEEE Transactions on Wireless Communications [C]. IEEE, 2021, 20(01):683–694.
32. Wu D, Sun X, Ansari N. An FSO-based drone assisted mobile access network for emergency communications [J]. IEEE Transactions on Network Science and Engineering, 2019, 7(3): 1597–1606.
33. Hossein Safi, Akbar Dargahi, Julian Cheng. Spatial Beam Tracking and Data Detection for an FSO Link to a UAV in the Presence of Hovering Fluctuations [J]. Signal Processing, 2019.
34. Safi H, Dargahi A, Cheng J. Beam tracking for UAV-assisted FSO links with a four-quadrant detector [J]. IEEE Communications Letters, 2021, 25(12):3908–3912.
35. Usman M, Althunibat S, Qaraqe K. Mobility dependent hybrid RF/FSO backhaul in UAV assisted cellular networks [A]. IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks [C].Italy: IEEE. 2020:1–6.
36. Nirwan Ansari, Di Wu, Xiang Sun. FSO as backhaul and energizer for drone-assisted mobile access networks [J]. ICT Express,2020,6(2):139–144.
37. Janji S, Samorzewski A, Wasilewska M, et al. On the Placement and Sustainability of Drone FSO Backhaul Relays [J]. IEEE Wireless Communications Letters, 2022, 11(08):1723–1727.

38. Lee J H, Park J, Bennis M, et al. Integrating LEO satellites and multi-UAV reinforcement learning for hybrid FSO/RF non-terrestrial networks [A]. Transactions on Vehicular Technology [C]. IEEE, 2023, 72(03):3647–3662.
39. Singh D, Reddy C S, Swaminathan R. Hovering UAV-Based FSO Communications with DF Relaying: A Performance Analysis [A]. National Conference on Communications [C]. India: IEEE. 2023:1–6.
40. An Jianxin, He Xiaolei, Yang Qianyuan, et al. Application research of small UAV air-ground wireless laser communication [J]. Optical communication technology, 2017, 41(06):10–13.
41. Xu Lin, Shen Yong, An Jianxin, et al. Development of an UAV airborne laser communication equipment [J]. Optical communication technology, 2018, 42(10):60–62.
42. Jiang X, Wu Z, Yin Z, et al. Power and trajectory optimization for UAV-enabled amplify-and-forward relay networks [J]. IEEE Access, 2018, 6(01):48688–48696.
43. Yang Liang, Yuan Jinhai, Liu Xinxin, et al. On the performance of LAP-based multiple-hop RF/FSO systems [J], IEEE Transactions on Aerospace and Electronic Systems, 2019, 55(01): 499–505.
44. Yuan Jinhai. Performance analysis of UAV Optical Wireless Communication system based on relay cooperation [D]. Guangzhou, Guangdong: Guangdong University of Technology, 2019.
45. Gu Zhiqun. Research on topology optimization technology of mobile fronthaul/backhaul network based on FSO[D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
46. Hu Mingfei. Research on downlink data optical transmission mechanism of light and small UAV [D]. Changchun, Jilin: Changchun University of Science and Technology, 2020.
47. Lei Xiaobo. Performance study of NOMA-based UAV cooperative FSO/RF communication system [D]. Guangdong Guangzhou: Guangdong University of Technology, 2020.
48. Y. Ma, J.-Y. Wang, J.-B. Wang, et al. Outage Performance Analysis and Parameter optimization of Hovering UAV-Based FSO System [A]. IEEE International Conference on Communications [C]. Ireland: IEEE. 2020:1–6.
49. Wang J Y, Ma Y, Lu R, et al. Hovering UAV-based FSO communications: Channel modelling, performance analysis, and parameter optimization [J]. IEEE Journal on Selected Areas in Communications, 2021, 39(10):2946–2959.
50. Liang Hanli. Design and implementation of automatic tracking control system for airborne laser communication [D]. Xi'an, Shaanxi: Xi'an University of Technology, 2021.
51. Wang J, Xu G, Zhao Z, et al. Performance Analysis of Dual-Hop UAV-Assisted FSO Communication System with DF Protocol [A]. 14th International Conference on Wireless Communications and Signal Processing [C]. Nanjing: IEEE. 2022:900–904.
52. Qu L, Xu G, Zeng Z, et al. UAV-assisted RF/FSO relay system for space-air-ground integrated network: A performance analysis [J]. IEEE Transactions on Wireless Communications, 2022, 21(8): 6211–6225.
53. Lu Rongrong. Performance analysis and optimization of UAV-based FSO communication system [D]. Jiangsu Nanjing: Nanjing University of Posts and Telecommunications, 2022.
54. Long Xiaoqiong. Research on performance analysis of relay wireless communication system based on UAV [D]. Guangzhou, Guangdong: Guangdong University of Technology, 2020.
55. Xu Su, Zhang Ji, Diao Yanghua, et al. Performance Analysis of UAV Relay Communication System Based on Hybrid Optoelectronic Transmission [J]. Telecommunications Science, 2022, 38 (04):113–120.
56. Ma Yang. Research on key technologies of FSO communication based on UAV relay [D]. Nanjing, Jiangsu: Southeast University, 2021.
57. Ma Jinwei. Ground user scheduling algorithm based on ergodic capacity in satellite-ground integrated network [J]. Fire and command control, 2022, 47(02):93–97+102.
58. Tang Hanling, Li Yongjun, Li Yi, et al. Performance analysis of UAV oblique laser communication based on MIMO [J]. Application optics, 2022, 43(05):1030–1049.
59. Li Chunruo, Yuan Junying. Dynamic switching algorithm between FSO and RF based on rate matching [J]. Journal of Missiles and Guidance, 2023, 43(02):32–50.

60. J. Wang, J. Wang, M. Chen, et al. Outage Analysis for Relay-Aided Free-Space Optical Communications Over Turbulence Channels with Nonzero Boresight Pointing Errors [J]. IEEE Photonics Journal. 2014, 12(6):1–15.
61. Ke Xizheng, Wu Jiali, Yang Shangjun. Research progress and prospect of atmospheric turbulence for Optical Wireless Communication [J]. Journal of Radio Wave Science, 2021, 36 (03): 323–339.
62. J. Park, E. Lee, C. Chae, et al. Impact of pointing errors on the performance of coherent free-space optical systems [J]. IEEE Photonics Technology Letters. 2016, 32(28):181–184.
63. S. Huang and M. Safari. Free-space optical communication impaired by angular fluctuations [J]. IEEE Transactions on Wireless Communications. 2017, 25(16):7475–7487.
64. Meng Lixin. Research on acquisition and tracking technology in airborne laser communication [D]. Changchun, Jilin: Jilin University, 2014.
65. Yan Lusheng, Wang Feng, Wu Wei, et al. The development status and key technologies of UAV laser communication payload [J]. Progress in laser and optoelectronics, 2016, 53(08):40–48.
66. Yang Shangjun, Ke Xizheng, Wu Jiali, et al. Fast alignment of Optical Wireless Communication using two-dimensional mirrors [J]. China Laser, 2022, 49(11):101–114.
67. Su Shaojing, Tong Xiaozhong, Wei Junyu, et al. Development status and trend of UAV swarm communication payload technology based on free space light [J]. Journal of National University of Defense Technology, 2021, 43(04):118–127.
68. Shen Y, Feng Y, Xu L, et al. Fast lift-off base station UAV airborne laser transmission technology [J]. Optical Communication Technology, 2019, 43(05):32–35.

Research Progress on Optical Wireless Communication in Industrial Internets

50

Contents

Introduction 1866
Optical Wireless Communication and the Industrial Internet 1866
Related Applications of Optical Wireless Communication in Industrial Scenarios 1868
Data Transmission 1868
Indoor Positioning 1868
Environmental Monitoring 1870
Foreign and Domestic Research Progress 1870
Key Technologies of Optical Wireless Communication in the Industrial Internet 1876
Visible Light Channel Model 1876
Multiple-Input Multiple-Output Technology 1878
Coding and Modulation Techniques 1880
Adjacent-Channel Interference Suppression Technology 1881
Multiple-Access Technology 1882
Heterogeneous RF/VLC Technology 1883
Optical Wireless Communication Problems in the Industrial Internet 1885
Long Transmission Range 1885
Indoor Attenuation 1885
Multipath Reflections 1885
Signal Loss and Blocking 1886
Theoretical Problems of Heterogeneous RF/VLC Transmissions 1886
Summary 1886
Device Positioning and SIGNAL Transmission 1886
Integrated Wearable Security Devices 1887
Underwater Industrial IoT 1887
Augmented-Reality Technology 1887
References 1887

Abstract

Optical wireless communication is an important means of modern communication, with rich spectrum resources, anti-electromagnetic interference, and other advantages, can be used as an important supplement to traditional radio frequency technology, and is expected to provide an important technical driving force for future

© Springer Nature Singapore Pte Ltd. 2024

X. Ke, *Handbook of Optical Wireless Communication*,
https://doi.org/10.1007/978-981-97-1522-0_50

industrial manufacturing. This chapter firstly introduces the relevant applications of optical wireless communication under the background of industrial Internet; summarizes the research status at home and abroad in recent years; summarizes the typical research progress; expounds the channel model and key technologies of optical wireless communication system; and introduces the RF/visible light heterogeneous technology used to ensure the reliable transmission of the uplink and downlink of the communication system. Finally, the chapter summarizes the current problems faced by optical wireless communication in the industrial Internet, and looks forward to the future development trend, which can provide certain reference and reference for the future research and development of optical wireless in this field.

Keywords

Industrial Internet · Optical wireless communication · Visible light communication · Data transmission · Visible light positioning

Introduction

Optical wireless communication uses light as a carrier to transmit information in free space. It has the advantages of rich spectrum resources, strong security, anti-electromagnetic interference, free spectrum authorization, and can easily monitor the rapid access of intelligent devices to Internet platforms.

Since the twenty-first century, with the continuous innovation in industrial production and industrial communication services, the industrial Internet has been gradually developing toward information and networks [1]. Traditional wired communication has high deployment costs and low transmission rates. The industry has proposed introducing a range of wireless-communication technologies into factories for data transmission, such as radio frequency identification (RFID), wireless LAN, and wireless sensor devices. However, these technologies face the problems of signal fading and multipath interference [2].

Optical wireless communication has been widely studied as an important supplement to radio frequency (RF) technology. Optical wireless communication under the industrial Internet is a subject worthy of further study and has practical significance.

This chapter introduces the application of optical wireless communication to the industrial Internet and summarizes and discusses the key technologies and research progress at home and abroad. Finally, the challenges of optical wireless communication in the industrial Internet are analyzed, and its future is examined.

Optical Wireless Communication and the Industrial Internet

In recent years, the industrial Internet has become an important support for China's development of a manufacturing power 2025 strategy. To enhance the safe and reliable interconnection between production factors and reduce costs, there is an

urgent need to replace traditional wired communication with wireless communication. Today's common RF wireless communication is vulnerable to electromagnetic interference from factory equipment and its spectrum is limited.

Visible light communication (VLC) combines lighting and communication through a high-density layout of light-emitting diodes (LEDs) to achieve a wide communication coverage range of factory equipment; not only can it transmit data to equipment and support indoor equipment positioning, but its spectrum resources are rich, more than 10,000 times [3] the existing RF spectrum. Hence, it can become a strong supplement to today's wireless communication technology. VLC and RF communication features are compared, as shown in Table 50.1.

VLC is an optical wireless communication technology [4] that transmits information in free space using 380–780 nm light waves as information carriers. A block diagram of the VLC system is shown in Fig. 50.1. It is mainly composed of three parts: a transmission module, wireless channel, and receiving module.

The transmission module is responsible for encoding and modulating the data to be transmitted to the corresponding electrical signal. The electrical signal passes through the LED drive circuit and becomes an optical signal. The LED transmits a flashing signal, which the human eye cannot recognize, to the receiving end [5] through a wireless channel. The receiving end primarily uses a photodetector or

Table 50.1 Comparison of VLC and RF communication characteristics [3]

Characteristic	VLC	RF
Bandwidth	428 THz–750 THz	300 MHz–10 GHz
Electromagnetic radiation	No	Yes
Multipath interference	Low	High
Noise interference	Large	Small
Power dissipation	Less	Relatively more
Facility	LEDs, no additional infrastructure is needed	Hot spots, infrastructure must be increased
System complexity	Low	High
Security	High	Low
Flexibility	Low	Relatively high

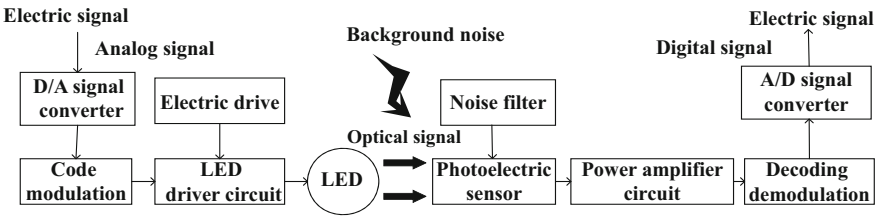


Fig. 50.1 VLC system block diagram [6]

image sensor to achieve the optical-to-electrical conversion, and the electrical signal is transmitted to the decoder and demodulation circuit to recover the original data.

Related Applications of Optical Wireless Communication in Industrial Scenarios

Data Transmission

Light fidelity (LiFi) extends the concept of VLC to enable two-way fully networked wireless communications [7]. Figure 50.2 presents an application block diagram for data transmissions in a factory. LED lights communicate through power-line communication (PLC). The upper computer loads the information on the power line and transmits it to intelligent devices in the factory through LED lights on the ceiling. The feedback information from the device is transmitted to the receiving end through the WiFi uplink and then transmitted to the host computer through PLC.

PLC can simplify the connection of equipment in a factory. There is no need to establish a new communication network as long as there are cables that can carry transmissions, to achieve low-cost and cost-effective broadband data delivery. Because LED lights require a power supply, power-line communication is a very attractive technology for LiFi backbone networks.

Indoor Positioning

With the development of intelligent factories, goods identification and positioning technologies are gradually being demanded. The wavelength of a light signal is short

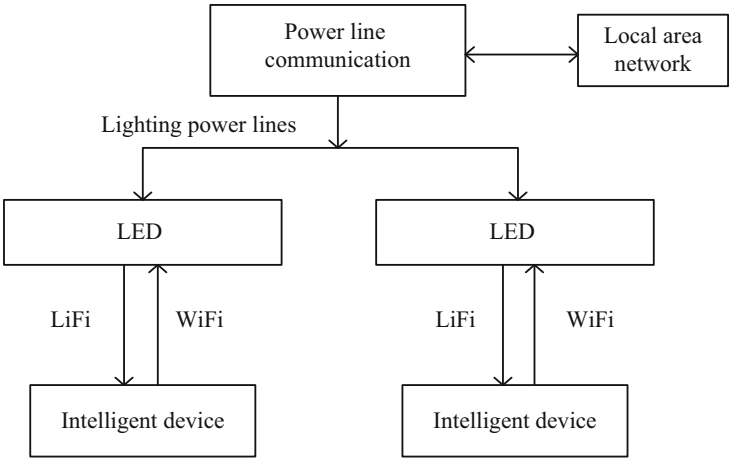


Fig. 50.2 Block diagram of factory data transmission [8]

and it propagates in a straight line, which makes visible light positioning (VLP) technology more accurate than RF positioning.

At the sending end of the optical signal line coding, the lamp is equipped with a modulator. Each LED is given a unique identifier that serves as a code word suitable for VLC transmission. The corresponding drive circuit opens the key control modulation mode to drive the LED lamp to achieve an electric-optical conversion. After the optical signal is received, it is decoded and processed; the position coordinates are then calculated using the corresponding positioning algorithm. The machine control module calculates [9] the direction information according to the positioning results, and inspects and navigates the factory equipment.

According to the receiving method, VLP can be divided into non-imaging-positioning technology based on a photodetector (PD) and imaging-positioning technology [10] based on an image sensor (IS). The positioning algorithms are classified as shown in Fig. 50.3.

Positioning technology with an IS as the receiver mainly uses an image-processing algorithm and optical-imaging principle to locate the measuring point. First, the camera captures the VLP signal and processes the received image. The position information [12] of the undetermined point is then calculated according to the optical-geometry relationship. Acceleration sensors commonly integrated into factory equipment can provide angle correction for images captured by this technology, thereby improving the positioning accuracy of the VLP.

Positioning technology with a PD as the receiver uses the PD's light sensitivity to receive visible light signals, combined with the time of arrival (TOA), time difference of arrival (TDOA), received signal strength (RSS), angle of arrival (AOA), scene analysis, and other positioning algorithms to achieve indoor positioning [13].

TOA and TDOA convert the time or time difference into a distance to measure it, and then use the trilateral positioning method to estimate the location. The RSS estimates the position according to the light signal strength from the anchor point to the LED. The AOA estimates the position by receiving visible light signals from multiple LEDs, according to the angle at which each LED signal reaches the PD receiver.

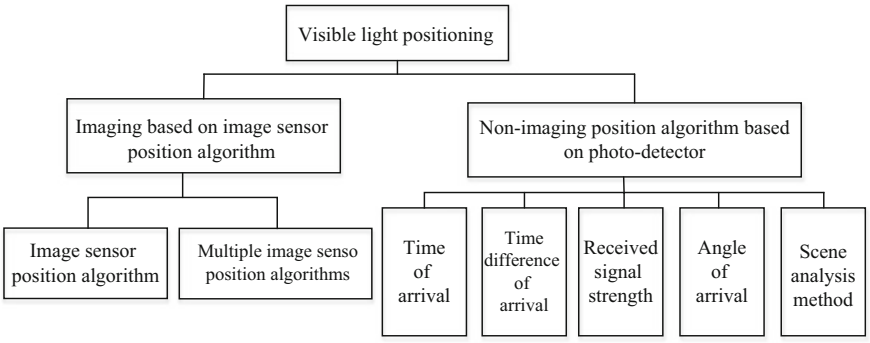


Fig. 50.3 VLP algorithm classification [11]

The scene analysis method first establishes the light intensity information of the environmental sampling point and records it in a database. After receiving the light intensity information, the receiver matches it with the database to estimate the location of the fixed point.

Environmental Monitoring

Some industrial sites such as chemical plants, mines, etc., are more sensitive to indoor environmental parameters and electromagnetic interference, such as gas concentration, indoor temperature, and air pressure in the current environment. In these cases, VLC can be combined with indoor environmental-monitoring technology, using the dual-use characteristics of LED communication lighting. The data collected by the sensor are sent to the background server through a VLC transceiver module to monitor the indoor environment in real time.

Figure 50.4 shows a block diagram of the VLC environmental-monitoring system in an industrial environment. The advanced RISC machine (ARM) processor is responsible for collecting environmental parameters, including temperature, humidity, air pressure, etc., and then sends the data information to the VLC sending module in the defined data frame format. The LED light is sent out in the form of visible light. The VLC receiving module converts the optical signal into a current signal, modifies it into a certain data frame format, and sends it to the server. The server is responsible for analyzing the received environmental data.

Foreign and Domestic Research Progress

As early as 1990, foreign scholars have proposed optical wireless communication (OWC) as an important means of indoor communication. At that time, there was great interest in the potential of free-space optical systems to provide portable data communication, especially with the congestion of the radio spectrum. Ultra-bandwidth frequency allocation was difficult to obtain, and microwave or millimeter-wave systems could not be used for higher bandwidth services, owing to technical limitations. Consequently, scholars in this field shifted their focus to

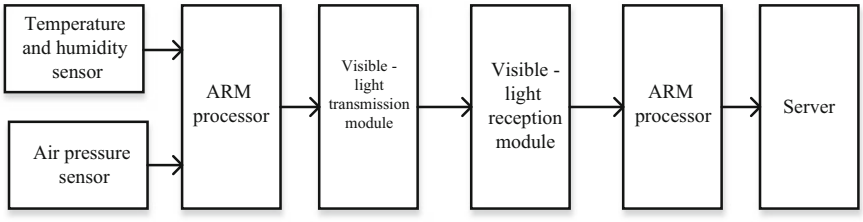


Fig. 50.4 Block diagram of a VLC environmental-monitoring system [14]

optical wireless communication technology, which was rich in spectral resources, security, and environmental protection.

In 2009, Paraskevopoulos et al. [15] proposed to apply OWC to short- and medium-distance industrial applications in the future. Orthogonal frequency-division multiplexing (OFDM) was proposed in combination with adaptive signal processing. Experiments showed that the transmission rate could be dynamically adjusted according to the current link quality. A reliable optical wireless link in the range of 100–200 Mbit/s achieved a high throughput.

In 2015, Goroshko et al. [16] studied the time-synchronization scheme of a VLC system based on OFDM. The system improved the transmission rate between machines in a factory, as well as the robustness and low-latency characteristics of the communication process, and could make future factory-manufacturing units more flexible.

In 2015, Shao et al. [17] proposed two heterogeneous WiFi–VLC systems. The first system used a one-way VLC channel as the downlink and retained the WiFi reverse channel as an uplink, which addressed the optical-uplink challenge. The second system used both WiFi and VLC parallel aggregation as a downlink. Studies showed that both hybrid systems had a high throughput, outperforming single WiFi systems, and that the aggregated system performed better when considering the blocking duration and distance between the access point and user device.

In 2017, Niu et al. [18] proposed a new intelligent lighting control system based on VLC, which combined with virtual reality technology can realize the information location of items in factories, the analysis of pipeline detection data, and the transmission speed is higher than the traditional intelligent lighting control system.

In 2018, Kim et al [19] proposed a VLC-based industrial Internet device management and data transmission scheme. Research showed that under this scheme, the VLC-packet exchange time was 590 ms, and the switching time between VLC transmitters was approximately 210 ms.

In 2018, Berenguer et al [20] studied channel measurement for robotic distributed 8×6 multi-input multi-output (MIMO) VLC systems in industrial manufacturing cells. The results showed that the link was available with a sufficient signal-to-noise ratio. However, because the direct-view link was blocked by the movement of the robot arm, the signal level suddenly dropped by 10–20 dB, which affected the link availability. An antenna diversity scheme was proposed to solve the link-blocking problem.

In 2018, Li Hao [21] proposed to use VLP as the indoor addressing mode of automatic guided vehicles in the factory, modulate the light source by using single-chip microcomputer cross-25 codes to improve transmission efficiency, use CMOS sensors to capture image information and process the image at the receiving end, and propose a dual-lamp positioning algorithm to realize automatic addressing of the car. The verification results showed that the system could satisfy the automatic addressing function; the positioning error was approximately 3.6 cm, and the positioning delay was approximately 37 ms.

In 2019, Berenguer et al. [22] proposed to increase the drive signal power in industrial OWC process by using a nonlinear equalizer, which can improve the

3.5 dB signal-to-noise ratio, and verified that the antenna diversity scheme in MIMO technology can significantly improve the reliability of the physical layer in the communication process.

In 2019, Almadani et al [23] proposed to combine VLP with UAVs to study the coverage area of UAVs in industrial environments. The results showed that when the divergence angle of the transmitter and the field-of-view angle of the receiver were set, the signal could be accurately received by the UAV outside a certain angle range.

In 2019, Lam et al. [24] proposed a new VLP ray-surface location method, which uses steerable lasers (rays) and LEDs to locate devices with active receivers. Studies showed that this technology could provide real-time tracking of devices within an operating area, add additional light sources for redundancy, scale to larger spaces, solve occlusion problems, provide capacity for more mobile devices, and be used as a low-cost positioning system for Industry 4.0.

In 2019, Liu et al. [25] proposed an indoor self-powered VLC system, which is able to collect energy from electromagnetic interference signals in factories to power VLC for IoT devices. An evaluation and analysis of the system proved its feasibility. It can be used as a supplementary technology in the hostile radio environment of the industrial Internet of Things.

In 2019, Yang et al. [26] proposed a heterogeneous RF/VLC industrial network architecture, which can simultaneously meet the high reliability, low latency, and high data transmission rate in industrial wireless networks, and proposed a new deep learning algorithm to realize intelligent resource management. Simulation results verified the effectiveness of the RF/VLC industrial network, and the proposed learning algorithm was superior to other existing algorithms.

In 2020, Bechadergue et al. [27] proposed that LiFi could provide a stable framework for the mass production of factory-specific chips based on the current market feedback of LiFi in the industrial field, laying a key foundation for the integration of LiFi technology into smart phones and other devices.

In 2020, Jungnickel et al. [28] proposed a new LiFi system concept for industrial wireless applications. It adopted a distributed multi-user MIMO architecture, which provided sufficient network capacity, moved seamlessly in large indoor areas, achieved reliable low-latency communication, and could be centralized into 5G core networks, such as WiFi.

In 2020, Muller [29] proposed a system architecture that combined 5G and LiFi in intelligent manufacturing scenarios using the industrial Internet. Research showed that this system provided reliable network communication and indoor positioning, making the integration of mobile devices in factories more flexible.

In 2020, Alturjman et al. [30] proposed to use subtle changes in the radiation intensity of LED lights to generate binary data streams, which are sensed by photodiodes and converted into electrical signals for wireless control of robots in factories, overcoming the need for wires. The results showed that the system improved the throughput and reliability for both direct and nondirect viewing links.

In 2020, Sun [31] introduced a LiFI-based IoT device access point architecture, which can meet the application scenarios of most IoT devices and realize energy

independence, providing a new idea for the future research of LiFi technology in the direction of IoT.

In 2021, Zhang et al. [32] proposed the diffuse reflection focusing technology and introduced the fast focusing transmission matrix algorithm, which was experimentally proved on the diffuse 25 cm free-space link. The experiment proved that a 40-Gbit/s non-line of sight beam could guide optical wireless links for communication, and the achieved function showed great potential in the Industry 4.0.

In 2021, Kouhini et al. [33] proposed a LiFi indoor positioning system based on time-of-flight measurements between multiple optical front ends deployed on the ceiling and devices moving within overlapping coverage areas. Studies showed that the system conformed to the recommended standard G.9991 of the International Telecommunication Standardization Department and achieved accurate indoor positioning with an average distance error of only a few centimeters. The LiFi positioning proposed in this study is a promising new feature that can be added to existing protocols to further enhance the functionality of intelligent lighting systems.

In 2021, Daniel et al. [34] conducted experimental research on the harsh environment of optical wireless channel transmission in factory conditions, proposed a new method to calculate the single spectral emission of superposition measurements, and built a model to quantify the impact of superposition measurements on communication. Research showed that the interference of the environment and artificial light sources affected the spectral and electrical-spectrum distributions. Therefore, the optical wireless communication system should be designed according to the actual factory environment.

In 2021, Koonen et al. [35] realized single wireless connection without interference of densely spaced devices in Industry 4.0 environment through narrow two-dimensional steerable infrared beams, capable of realizing self-calibration positioning of devices, and proposed a low-complexity wideband wide-field optical wireless receiver based on a photodiode matrix.

In 2021, Kang et al. [36] will design and optimize the optimal rotation angle receiver and aperture array receiver for complex dynamic scenes in the industrial Internet of Things to reduce noise interference and interference between adjacent channels during VLC transmission. Research showed that the design enabled robust data transmissions and improved the anti-interference ability.

In 2022, Ma et al. [37] evaluated the LiFi positioning system based on ITU-T G.9991 OWC standard in the factory environment, and the results showed that the three-dimensional positioning errors of the system were 7, 6 and 3 cm, respectively, and were further improved to about 3, 5, and 4 cm by introducing correction factors. Their research method demonstrated the possibility of further developing LiFi systems for product positioning and navigation vehicles in the future industrial Internet.

In 2022, Salvi et al. [38] proposed an optical camera communication system inspired by the Aztec model that could be applied in industrial scenarios in the future. The signal-to-noise ratio and intensity distribution of the system were evaluated, and the results showed that the method was superior to traditional

two-dimensional code recognition methods, in terms of power consumption and bit-error rate.

In 2022, Kouhini [39] proposed a real-time communication system for an all-optical fixed-wireless LiFi link with distributed MIMO. The operation of this link under spatial diversity and spatial multiplexing was analyzed, and the research showed that dynamically switching between spatial diversity and spatial multiplexing was feasible.

In 2022, Costa et al. [40] used artificial intelligence based on machine learning to predict the trajectory and channel performance of OWC links. Two industrial scenarios were considered as examples to evaluate the proposed scheme. The results showed that the algorithm could optimize the link parameters, predict the VLC channel gain, and effectively improve the power and spectral efficiency of the VLC system in industrial scenarios.

In 2022, Reguera et al. [41] proposed a VLC system based on multi-pulse position modulation (MPPM), which allows dimming control and can be used in resource-constrained devices. The results showed that the system had an efficient data transfer rate and lower energy consumption. The proposed new codec algorithm for MPPM significantly reduced the running time and could easily be embedded in the LED drivers and receivers of VLC for large-scale industrial deployment.

In 2022, Xue et al. [42] proposed a hybrid integrated system combining multi-access edge computing and visible light communication to solve the problem of unbalanced resource utilization caused by users moving in the dense coverage of VLC. The results showed that the system had a lower user drop rate, improved the resource utilization, reduced the overall delay, and met the requirements of low delay and high capacity in today's industrial IoT.

In 2022, Xue et al. [43] proposed a no-boundary dense networking VLC system, which cancels the multiple access technology of transmitter in traditional industrial scenes to improve transmission efficiency, and the receiving end can adaptively realize anti-noise reception in nonoverlapping areas and simultaneous recovery of two transmitted signals in overlapping areas. The results showed that the system had a lower bit-error rate and wider communication coverage.

In 2022, Xu [44] proposed four optical OFDM schemes based on discrete Hartley transform LED index modulation and subcarrier index modulation to reduce hardware costs and computational complexity. The results showed that the four proposed schemes obtained better peak-to-mean ratios and error performance, providing a reference for introducing efficient optical OFDM system schemes in industrial Internet scenarios.

In 2022, Yang [45] proposed three time-synchronization algorithms for DC-biased optical OFDM systems applied to VLC. Research showed that the proposed algorithms had a higher time-synchronization performance, reduced the impact of time-synchronization errors on the system performance, better complemented RF technology in industrial scenarios, and had a higher spectral efficiency and lower algorithm complexity.

In 2022, Ozyurt et al. [46] proposed an industrial LiFi network technology combining device-to-device communication and coordinated multi-point joint

transmission, and launched an analytical model of the coverage area of inter-device communication according to half illuminance and half angle. The simulation results agreed with the proposed theoretical model, which was valuable for practical industrial LiFi design and development.

In 2022, Yang [47] combined VLC and PLC technology, used the power-line transmission data as the uplink, selected their own adaptive OFDM technology on the two transmission channels, and built a PLC-VLC cascade communication system platform for testing. The results showed that the system could achieve stable upstream and downstream communication, and the transmission bit-error rate reached 10^{-4}, which opened a new idea for upstream and downstream communication links in industrial settings.

In 2023, Nguyen et al. [48] proposed a MIMO modulation system based on time-domain open key control modulation and light-emitting diode array. The results showed that compared with traditional region of interest signal detection, the deep learning for detection and tracking of the system has the advantage of low bit-error rate over a long communication distance, and can be widely used in smart factories.

In 2023, Wang et al. [49], based on the intelligent lighting system, proposed combining the visible light system with virtual reality technology to achieve positioning and automatic detection to meet the information requirements of smart factories.

In 2023, Du et al. [50] proposed a visible light positioning LED region recognition framework based on CMOS image sensor, which can realize mobile terminal positioning from the perspective of recognition without modulating the LED. Experiments showed that the model was highly robust and could be applied to various types of LED lights, which play an important role in the positioning of mobile terminals in industrial environments.

In 2023, Zhou et al. [51] suppressed cross talk in full-duplex VLC by reducing beam angle and field angle, power isolation, and ground plane segmentation. The results showed that, combined with the existing communication network, a full-duplex real-time VLC system based on Ethernet was realized, and the system could also be used as an optical wireless access system for the industrial Internet.

In 2023, Aboagye et al. [52] introduce the design of VLC system based on reconfigurable intelligent surface and integrate reconfigurable intelligence surface (RIS) with other emerging VLC technologies, so as to solve the problem of direct-view blocking in VLC system. Research showed that by selecting an appropriate RIS technology to adapt to VLC systems, the complexity of the RIS technology could be effectively reduced, and the data transmission rate could be improved. This design can provide useful guidance for wireless communication networks in the industrial Internet.

In 2023, Ozyurt et al. [53] analyzed the key performance indicators of device-to-device (D2D) communication in ultra-dense industrial LiFi networks, and deduced the coverage model of LiFi for D2D communication range and the expression of D2D residence time. Research showed that densely deployed access points led to a reduced D2D residence time and increased transition rate, which are of great value in practical industrial LiFi design.

Key Technologies of Optical Wireless Communication in the Industrial Internet

Visible Light Channel Model

The VLC system adopts a Lambert-type white LED at the transmitting end and a photoelectric sensor at the receiving end. The VLC linear baseband-transmission model is shown in Fig. 50.5.

The photocurrent signal received by the receiving end is represented as

$$Y(t) = R \cdot X(t) \otimes h(t) + N(t) \tag{50.1}$$

where $\otimes$ is a convolution, R is the photoelectric conversion efficiency, $X(t)$ is the transmitted light power of the LED, $h(t)$ is the impulse response of the channel, and $N(t)$ is the additive Gaussian white noise in the channel.

In a closed indoor VLC environment, an LED optical signal can be divided into two propagation paths to reach the receiving end. The path directly to the receiving end is called the line of sight (LOS), and the path to the receiving end after reflecting off an object in the propagation process is called the non-line of sight (NLOS) [55]. Owing to the presence of solid particles, such as smoke and dust, in some plants it is necessary to consider the effect of the particle-scattering characteristics on the light propagation, according to the actual situation in the channel analysis. The VLC link modes are shown in Fig. 50.6.

In the LOS link, the transmitting and receiving ends are aligned and have a high power utilization rate. However, if there are obstacles between the LED light source and the receiving end, the signal transmission will be hindered or communication will be interrupted. In the NLOS link, the receiving angle is large, and the requirements for the directions of the transmitting and receiving ends are low.

However, there is a difference in the time required for the reflected light signals of different objects to reach the receiving end, which may lead to intercode cross talk, thereby reducing the accuracy of the VLC system. Therefore, in actual factory applications, the two link modes should be selected according to the different requirements.

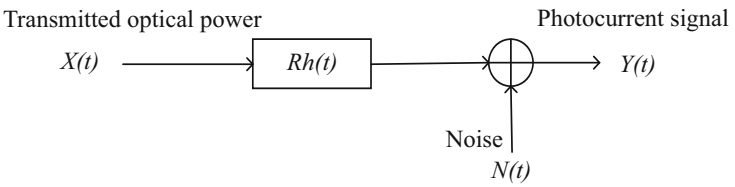


Fig. 50.5 Linear baseband-transmission model for VLC [54]

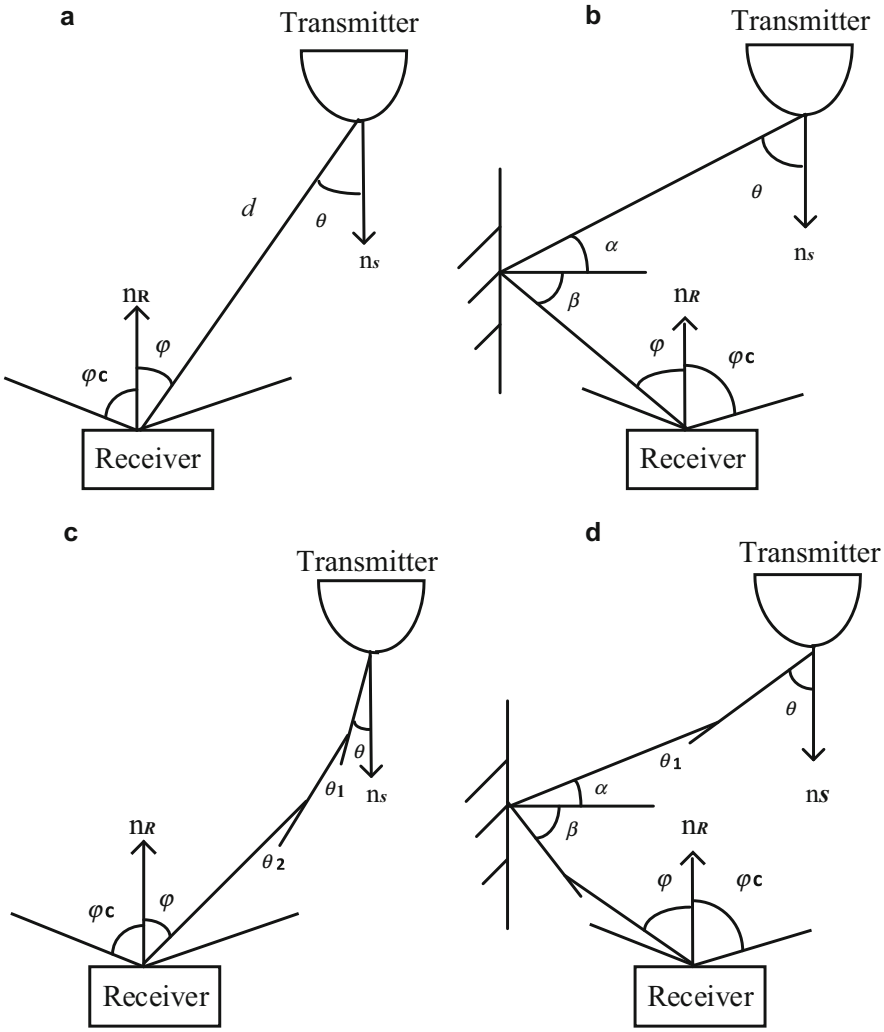


Fig. 50.6 VLC link modes [56]. (**a**) Line of sight; (**b**) non-line of sight; (**c**) LOS (with scattering); (**d**) NLOS (with scattering)

The power formula of the receiving device is

$$P_r = H(0) = H(0) \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X(t) d(t) \tag{50.2}$$

where P_r denotes the mean value of the transmitted power; $H(0)$ denotes the DC channel gain; and $H_d(0)$ and $dH_r(0)$ are successively summed in the LOS and NLOS links, respectively. The formula is as follows:

$$H_d(0) = \begin{cases} 0 & ,\varphi > \varphi_c \\ \dfrac{(m+1)A}{2\pi d^2} \cos^m(\theta) T(\varphi) g(\varphi) \cos(\varphi), & 0 \le \varphi \le \varphi_c \end{cases} \tag{50.3}$$

$$dH_r(0) = \begin{cases} 0 & ,\varphi > \varphi_c \\ \dfrac{(\mathrm{m}+1)A}{2(\pi d_1 d_2)^2} \cdot dA_{\mathrm{wall}} \cdot \cos^{\mathrm{m}}(\theta) \cos(\alpha) \cos(\varphi) T(\varphi) g(\varphi) \cos(\varphi), & 0 \le \varphi \le \varphi_c \end{cases} \tag{50.4}$$

In Eqs. (50.3) and (50.4), A represents the receiving area of the sensor; θ represents the transmitting angle; φ represents the receiving angle; and φ_c represents the receiving range of the receiver. When the incident light exceeds the receiver range, the optical signal will not be received. m is the Lambert radiation coefficient; d represents the link distance; d_1 represents the distance from the LED sending array to the first reflection point; d_2 represents the distance between the first reflection point and the receiver; and dA_{wall} is the unit area of the reflecting surface. α and β are, respectively, the incident angle of light to the reflection point and the radiation angle of the reflection point to the receiver. $T(\varphi)$ represents the filter gain and $g(\varphi)$ represents the intensity of the concentrated energy of the receiver.

The impulse response $h(t)$ is shown in Eq. (50.5):

$$h(t) = \frac{(\mathrm{m}+1)A}{2\pi d^2} \cos^{\mathrm{m}}(\theta) T(\varphi) g(\varphi) \cos(\varphi) \mathrm{rect}\left(\frac{\varphi}{\varphi_c}\right) \delta\left(t - \frac{d}{c}\right) \tag{50.5}$$

Multiple-Input Multiple-Output Technology

With the rapid increase in the number of devices on the industrial Internet, it is necessary to establish powerful communication networks to ensure reliable communication between devices. To support large-scale connections [57–59] in industrial networks, multiple-input multiple-output (MIMO) technology has been introduced into the VLC transmission process. This technology can effectively use space resources, realize multi-antenna parallel data transmissions, and greatly improve the transmission capacity [60] of the system without increasing the system transmission bandwidth.

Visible MIMO technology involves setting up a number of antennas at both ends of the transceiver. The transmitter is a multichannel serial signal modulated and encoded using visible-channel parallel transmissions. The receiving end of the signal demodulates, decodes, and restores the original signal [61]. A schematic of the visible MIMO system is shown in Fig. 50.7.

Visible MIMO technologies can be divided into two kinds [63]: spatial diversity and spatial multiplexing.

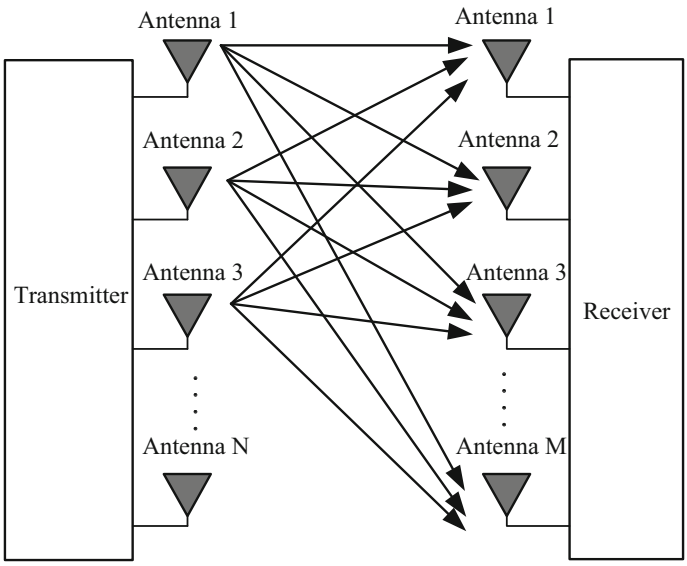


Fig. 50.7 Schematic diagram of a visible MIMO system [62]

Spatial Diversity

Spatial diversity uses multiple transmitting antennas to transmit the same data through different communication channels and multiple receiving antennas to receive information [64] at the receiving end. Because the signal attenuation differs in different channels, the receiver combines multiple signals, which can effectively reduce the bit-error rate of the system, weaken the influence of multipath fading in the channel, and improve the reliability of communication transmissions without increasing the transmission power and bandwidth.

Spatial Multiplexing

Spatial multiplexing uses multiple antennas at the sending end to transmit independent signals simultaneously. To achieve spatial multiplexing, it is necessary to ensure that the distance between the sending and receiving antennas is greater than the correlation distance to ensure that the subchannels at both ends of the transceiver are independent and unrelated. However, the adjacent channels may interfere with each other; therefore, it is necessary to adopt precoding technology to preprocess the signal at the transmitting end, or adopt anti-jamming technology, such as the zero-forcing, minimum mean-squared error, and maximum-likelihood algorithms, at the receiving end to reduce the bit-error rate [65] of the system.

Although spatial diversity is similar to spatial multiplexing, the two methods divide the data into several information streams and send them out through different transmitting antennas. However, the gain of spatial multiplexing is not as high as that of spatial diversity. The information flow of spatial diversity is used to send redundant information and improve the reliability of the communication

transmission, whereas spatial multiplexing can increase the transmission rate and improve the effectiveness of the communication transmission by sending different information flows.

MIMO technology is a compromise between the bit-error rate and data transmission rate of the system. While improving the spatial diversity gain, the spatial multiplexing gain is reduced [63], so it is necessary to select the appropriate technology according to the actual communication purpose in the factory.

Coding and Modulation Techniques

VLC systems usually adopt an intensity modulation/direct detection mode to transmit signals. The encoding mode mostly adopts binary open-key control or Manchester coding to improve the transmission speed. It also uses similar optical-group encoding forms, such as pulse position modulation; however, the system has higher requirements for clock synchronization in the decoding process. Binary on–off keying is relatively simple; however, the transmission speed is slow and the time-slot error rate is high. The Manchester coding bit-error rate is low, but the frequency-band bandwidth requirements are high, and the available modulation bandwidth of LEDs is very limited [66].

In VLC system modulation technology, the commonly used single-carrier modulation scheme causes a series of problems [67] such as LED front-end nonlinear signal distortion, and the data transmission rate is high in the industrial environment. In order to achieve high-speed transmission of information under the limited channel bandwidth, researchers at home and abroad consider introducing more efficient multi-carrier modulation technology into the industrial VLC systems.

Orthogonal frequency-division multiplexing (OFDM) technology has the advantages of high spectral efficiency and strong bandwidth expansibility, which can effectively reduce the frequency-selective fading caused by the indoor multipath effect. The basic concept is to divide the entire communication channel into multiple smaller subchannels, divide serial data into multiple parallel data streams, and transmit information on each subchannel after modulation.

The signal bandwidth on the subchannel is smaller than the relevant bandwidth of the entire channel, which can resist delay expansion and overcome the interference between signal symbols. The optical OFDM block diagram is shown in Fig. 50.8.

After the input data stream undergoes constellation modulation and series/parallel conversions, the frequency-domain signal is converted into a time-domain signal through an N-point Fourier inversion. To ensure the orthogonality between subchannels, a cyclic prefix is inserted, and a continuous analog signal is obtained after parallel/series and digital-mode conversions. Then, a DC bias is added to ensure that the signal is a positive real-value signal, and it is finally transmitted by the LED light source. After receiving the signal, the receiving end performs the inverse process of the transmitting end to restore the original data [68].

In the communication transmission process, each subcarrier adopts the same transmit power and simultaneously carries out the same modulation mode, which

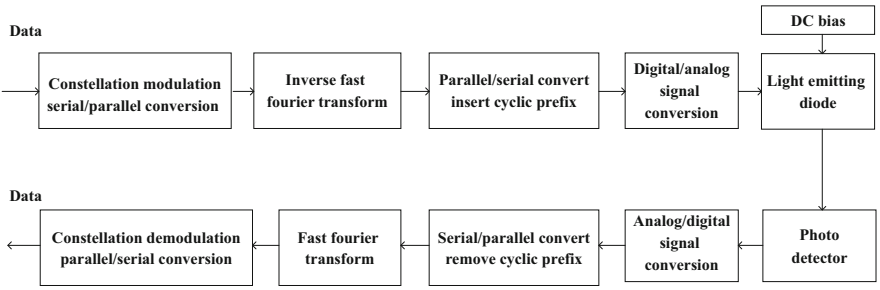


Fig. 50.8 Optical OFDM block diagram

may cause an increase in the bit-error rate among the subchannels and affect the system performance.

In 2009, Turkish Paraskevopoulos [15] proposed to introduce adaptive technology into an OFDM system. It achieved reliable communication links in the range of 100–200 Mbit/s and reduced the multipath effect of the system. Adaptive OFDM modulation can reasonably adjust the modulation mode and power, according to the transmission situation of each subchannel, to realize high system flexibility, maximize the utilization of channel resources, and improve the transmission efficiency [69] of the system.

Adjacent-Channel Interference Suppression Technology

Co-channel interference (CCI) [3] will be generated when the same transmission frequency is used. CCI schematic diagram is shown in Fig. 50.9. The emission mode of the LED light source has strong directivity, and most of the optical power radiated by the light source is located in the coverage area of the access point. Therefore, CCI will occur in the area between the edges of the coverage unit and the adjacent access point, where the optical emissions overlap. For high-density spatial multiplexing, CCI significantly affects the accuracy [70] of the communication transmission. The problem of CCI in adjacent units in visible optical communications can be effectively suppressed by adopting frequency-multiplexing technology.

A VLC system with multiple LED arrays is known as a cellular optical network. The basic idea of frequency-multiplexing technology is to divide the entire optical network into several cell clusters, and cells with the same frequency are separated by a certain distance. The number of units that use the same frequency in a cluster is called the frequency-multiplexing factor [71].

A schematic diagram of frequency multiplexing is shown in Fig. 50.10, where (a) is unified-frequency multiplexing and (b) is partial-frequency multiplexing with a frequency-multiplexing factor of two. All cells in unified-frequency multiplexing use the same frequency band, and all cells in partial-frequency multiplexing with a frequency-multiplexing factor of two share two frequency bands. Interference occurs

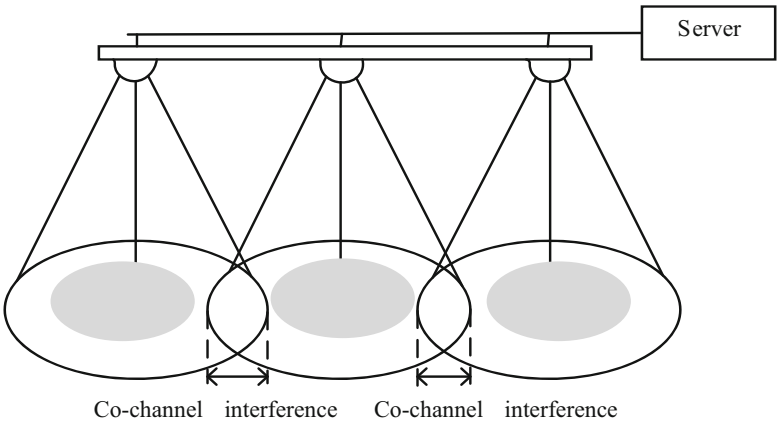


Fig. 50.9 Co-channel interference diagram

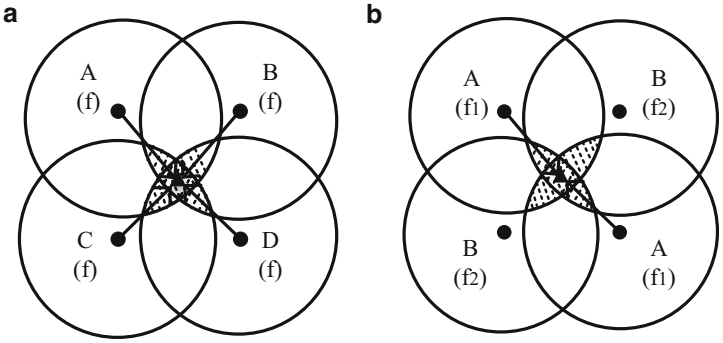


Fig. 50.10 Frequency-reuse diagram [71]. (**a**) Unified-frequency multiplexing. (**b**) Partial-frequency multiplexing

only in the overlapping areas of adjacent cells at the same frequency, thereby suppressing the CCI.

Multiple-Access Technology

In a factory, when multiple devices communicate simultaneously using the same resource (such as a shared channel), information collisions can result in communication failures. Multiple-access technology means that multiple devices can access the same communication network simultaneously. They use a specific channel-multiplexing technology in the same transmission channel, in which multiple devices can independently and reliably receive and demodulate data to share resources and improve channel utilization.

Table 50.2 Comparison of common multiple-access technologies

Multiple-access method	Principle	Advantages	Disadvantages
Time division	Different devices transmit information in different time slots	The average power efficiency is high and the realization difficulty is low	Requires equipment synchronization and has high overhead
Frequency division	Different devices transmit information in different frequency bands	Multiple signals are transmitted simultaneously without interfering with each other	The more frequency bands are divided, the slower the transmission rate on a single channel
Code division	Orthogonal codeword sequences are used to distinguish different terminal signals	Bandwidth communication has strong anti-fading and anti-jamming ability	When there are many users, the code word is too long
Space division	Free space is divided into different subchannels to transmit data	High spectrum utilization	High precoding complexity, interference from adjacent channels
Nonorthogonal	Different transmit powers are assigned to terminals according to the channel quality	Multiple signals are transmitted simultaneously, with a high transmission rate	High complexity

Currently, a variety of multiple-access technologies [72] have been developed for indoor VLC systems. Their principles, advantages, and disadvantages are listed in Table 50.2.

Heterogeneous RF/VLC Technology

Obstacles between the receiving and sending ends in a factory or changes in the position of the receiving equipment may lead to a reduction in VLC signal quality or communication interruptions. An effective method to address this challenge is to combine the advantages of RF wireless communication with those of VLC high-speed communication to form a heterogeneous RF/VLC network [73].

According to current research, the uplink of a communication system typically adopts RF wireless communication technology, and the downlink adopts heterogeneous RF/VLC technology. System models can be roughly divided into three types: double-hop, independent, and parallel-transmission [74].

Double-Hop RF/VLC System

As shown in Fig. 50.11, the double-hop RF/VLC system divides the link transmissions into two stages. The first stage is set as the VLC link, and the second stage is set as the RF link. Between the two hops, relay nodes, intelligent reflectors, and

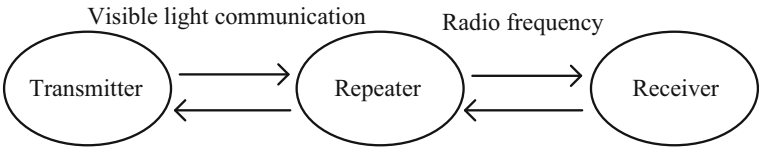


Fig. 50.11 Double-hop heterogeneous RF/VLC system

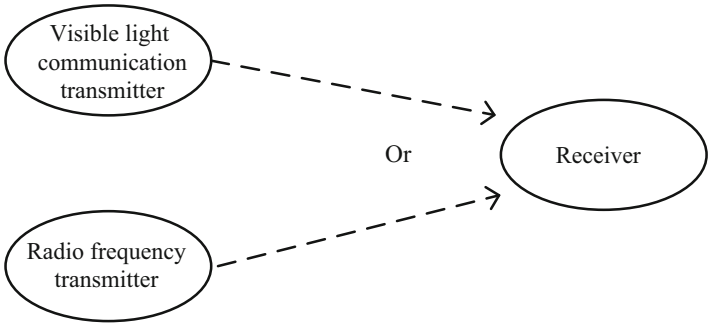


Fig. 50.12 Independent RF/VLC system

other devices are added to convert the VLC transmission signals into signals suitable for transmission in the RF link.

Conversely, the first hop can be set as an RF link and the second hop as a VLC link. This system can be applied to the receiving end but does not have the ability to receive VLC or RF system transmission signals; the direct link between the receiving and sending ends is unreachable.

Independent RF/VLC System

As shown in Fig. 50.12, an independent RF/VLC system implies that the receiver or central unit set by the system decides whether to use an independent VLC link or an independent RF link for information transmission. Typically, the receiver in the system does not have the ability to aggregate VLC and RF information.

The receiving end can freely choose the best network, according to the received signal-to-noise ratio, application requirements, device mobility, and other parameters, while the central unit set by the system can choose the network to provide services, according to the load balance and fairness of the network.

Parallel-Transmission RF/VLC System

As shown in Fig. 50.13, a parallel-transmission RF/VLC system implies that the RF and VLC systems simultaneously transmit information to the receiving end, and the receiving end uses aggregation technology to integrate the sent information. The real-time transmission performance of the heterogeneous system is good. When one link does not work, the other sends and receives information normally. However, the equipment complexity of this system is relatively high and difficult to realize.

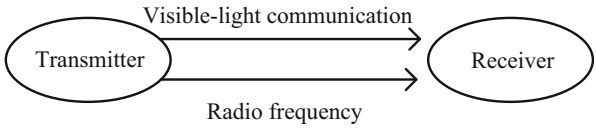


Fig. 50.13 Parallel-transmission heterogeneous RF/VLC system

Optical Wireless Communication Problems in the Industrial Internet

Although research on optical wireless communication technology in the industrial Internet has achieved some results, the actual application of these technologies in the industry is immature, and many challenges remain before widespread deployment. These include a long transmission range, indoor attenuation, multipath reflection, signal loss and blocking, and heterogeneous RF/VLC transmission theory.

Long Transmission Range

In larger places, such as warehouses and factories, the ceiling height leads to a longer link range, which causes multipath dispersion, resulting in higher path losses and lower data rates. In addition, the more light sources on the ceiling, the higher the level of interference between them, which affects the communication system performance [75].

Indoor Attenuation

The transmission medium for indoor applications is usually fresh air. Complex industrial environments may contain impurities, such as oil vapor, water mist, and industrial smoke, which affect light-wave transmissions and weaken the signal power [76].

Multipath Reflections

Highly reflective surfaces, such as metal devices and equipment, in industrial environments cause the multipath reflection of optical signals. The intersymbol interference generated by the multipath effect affects the quality of optical-signal transmissions. According to current research, OFDM can be used to improve the performance of VLC links, therefore different variants of OFDM have been introduced, such as DC-biased optical OFDM, asymmetric limiting optical OFDM, pulse–amplitude discrete multitone systems, etc.

Signal Loss and Blocking

VLC relies heavily on direct-vision propagation; therefore, there must be a clear and uninterrupted path between the sending and receiving ends. The random motion of a robot arm in a factory is likely to block light-wave propagation and cause signal loss. Foreign researchers have proposed an antenna diversity model [20] that places the receiving end near the manufacturing unit to ensure that there is always a direct-view link; the results showed that even if the robot arm moved a few centimeters, a signal attenuation of up to 20 dB could occur.

Theoretical Problems of Heterogeneous RF/VLC Transmissions

For heterogeneous VLC/RF networks, most existing studies assume that the input signal of the VLC is a Gaussian signal, but the visible signal is bounded and nonnegative; therefore, the traditional Shannon formula is not suitable for deriving the transmission rate. The closed-form expression of the transmission rate of heterogeneous RF/VLC systems and the optimal-transmission mechanism require further investigation.

In addition, the current design of optical wireless communication in the industrial-Internet environment is still limited to Lambert beam emitters, which do not fully use the non-Lambert beam-space characteristics of today's commercial LED light sources. Owing to the complexity of industrial environments, research in related fields must be evaluated and characterized to determine whether it meets the industrial-automation requirements to be placed into factories.

Summary

In recent years, researchers at home and abroad have demonstrated the industrial application potential of optical wireless communication through numerous experiments. Future research on optical wireless technology in the industrial field will focus on improving bandwidth, increasing transmission distance, and automatic direction alignment. Currently, it has the following potential applications in the industrial Internet environment.

Device Positioning and SIGNAL Transmission

Optical wireless communication is used for device positioning and signal transmission. A mobile vehicle in a factory can exchange data with the LED light and position itself using the power intensity emitted by the LED. The central controller can accurately locate the position of the mobile vehicle through the feedback signal conveyed by the LED. The device installed with the LED can transmit data

information and request the delivery of raw materials for production. A drone will respond immediately and execute the task after receiving the corresponding signal.

Integrated Wearable Security Devices

In some high-risk industrial places, workers must wear wearable safety devices in the form of helmets to cope with emergencies. VLC transceivers can be integrated into wearable safety devices to capture warnings emitted by the LEDs and broadcast them to other employees to improve workplace safety.

Underwater Industrial IoT

In marine environments, the underwater industrial Internet is used to transmit data throughout the ocean, enabling underwater sensor systems to communicate with each other and transmit data to ground networks. Because radio frequency signals cannot penetrate water, the data rate of acoustic signals in the transmission process is low and the cost is high; therefore, optical wireless communication is considered one of the communication technologies for the underwater industrial Internet.

Augmented-Reality Technology

Augmented reality is a technology that complements real-world experiences with the help of virtual information to provide guidance for machinery operation in factories. Augmented reality can be combined with wireless light to virtually label data for a machine, and employees can use dedicated devices to access this information, making it easier to work.

References

1. Wang Min. Research on wireless communication technology for Industrial Internet [J]. Changjiang information communication, 2022, 35(10): 227–235.
2. Ding Jupeng, Yi Zhiling, Zhao Kai, et al. Research progress of industrial wireless optical technologies [J]. Optical Communication Technology, 2022, 46(2): 35–40.
3. Zhang Siyu. Research on spectrum resource allocation and access scheme of LiFi network [D]. Dalian: Dalian Maritime University, 2019.
4. Huang H M, Lan T X, Huang C, et al. Visible-light communication using high-power LED panel-lamp and low-complexity Mosfet circuit [J]. Optical and Quantum Electronics, 2021, 53(1): 1–17.
5. Li Zheng, Wang Feigang, Liang Jingyuan, et al. The Principle and Research Progress of LED Visible Light Communication [J]. China Illuminating Engineering Journal, 2023, 34(1): 29–40 +44.
6. Zheng Han. Research on indoor UAV location method based on visible light [D]. Chongqing: Chongqing University of Posts and Telecommunications, 2021.

7. Zhang Liping. Overview of LiFi technology developments [J]. Modern Science and Technology of Telecommunications, 2017, 47(2): 42–48+55.
8. Yang Shirong. Discussion on application of LiFi technology in ship integrated monitoring [J]. Automation and Instrumentation, 2022, 37(11): 14–16+43.
9. Luo J H, Fan L Y, Li H S. Indoor Positioning Systems Based on Visible Light Communication: State of the Art [J]. IEEE Communications Surveys and Tutorials, 2017, 19(4): 1–4.
10. Li Ning. Research and system implementation of indoor positioning algorithm [D]. Beijing: Beijing University of Posts and Telecommunications, 2018.
11. Guo Tao. Research on VLC channel modeling and localization for indoor irregular scenes [D]. Baotou: Inner Mongolia University of Science and Technology, 2022.
12. Wang Y Q, Gong Y K, Shi Z F. Research on the Collinear Equation Model of Visual Positioning Based on Visible Light Communication [J]. MATEC Web of Conferences, 2015, 22(4): 33–50.
13. Jiang J J, Guan W P, Chen Z N, et al. Indoor High-precise Three dimensional Positioning Algorithm Based on Visible Light Communication and Fingerprinting Using K-Means and Random Forest [J]. Optical Engineering, 2019, 58(1): 336–348.
14. Peng Kai. Research on portable wireless data transmission system based on visible light [D]. Xi'an: Xi'an Technological University, 2017.
15. Paraskevopoulos A, Vucic J, Voss S H, et al. Optical free-space communication systems in the Mbps to Gbps range suitable for industrial applications [C]// 2009 International Symp Parosium on Optomechatronic Technologies, September 21–23, 2009, Istanbul, Turkey. New York: IEEE, 2009: 377–382.
16. Goroshko K, Manolakis K, Grobe L, et al. Low-latency synchronization for OFDM-based visible light communication [C]// 2015 IEEE International Conference on Communication Workshop, June, 8–12, 2015, London, UK. New York: IEEE, 2015: 1327–1339.
17. Shao S H, Khreishah A, Ayyash M, et al. Design and Analysis of a Visible-Light-Communication Enhanced WiFi System [J]. Journal of Optical Communications and Networking, 2015, 7(10): 960–973.
18. NIU Yanfang, Sun Dedong, Gong Taorong, et al. Application of Visible Light Communication Technology in the Intelligent Industrial Lighting [J]. China Illuminating Engineering Journal, 2017, 28(5): 27–38.
19. Kim C M, Koh S J. Device Management and Data Transport in IoT Networks Based on Visible Light Communication [J]. Sensors, 2018, 18(8): 2741–2752.
20. Berenguer P W, Schulz D, Hilt J, et al. Optical wireless MIMO experiments in an industrial environment [J]. IEEE Journal on Selected Areas in Communications, 2018, 36(1): 185–193.
21. Li Hao. Design of indoor automatic positioning guidance system based on visible light positioning automatic guided vehicle [D]. Guangzhou: Jinan University, 2018.
22. Berenguer P W, Hellwig P, Schulz D, et al. Real-time optical wireless mobile communication with high physical layer reliability [J]. Journal of Lightwave Technology, 2019, 37(6): 1638–1646.
23. Almadani Y, Ijaz M, Rajbhandari S, et al. Dead-zones limitation in visible light positioning systems for unmanned aerial vehicles [C]// 2019 Eleventh International Conference on Ubiquitous and Future Networks, July, 2–5, 2019, Zagreb, Croatia. New York: IEEE, 2019: 419–421.
24. Lam E W, Little T D C. Visible Light Positioning for Location-Based Services in Industry 4.0 [C]// 2019 16th International Symposium on Wireless Communication Systems, August 27–30, 2019, Oulu, Finland. New York: IEEE, 2019: 345–350.
25. Liu X Y, Wei X T, Guo L, et al. Turning the Signal Interference Into Benefits: Towards Indoor Self-Powered Visible Light Communication for IoT Devices in Industrial Radio-Hostile Environments [J]. IEEE Access, 2019, 7(32): 1761–1772.
26. Yang H L, A Alphones, Zhong W D, et al. Learning-Based Energy-Efficient Resource Management by Heterogeneous RF/VLC for Ultra-Reliable Low-Latency Industrial IoT Networks [J]. IEEE Transactions on Industrial Informatics, 2020, 16(8): 5565–5576.

27. Béchadergue B, Azoulay B. An Industrial View on LiFi Challenges and Future [C]// 2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing, July, 20–22, 2020, Porto, Portugal. New York: IEEE, 2020: 1–6.
28. Jungnickel V, Berenguer P W, Mana S M, et al. LiFi for Industrial Wireless Applications [C]// 2020 Optical Fiber Communications Conference and Exhibition, March, 8–12, 2020, San Diego, CA, USA. New York: IEEE, 2020: 1–3.
29. Müller M, Behnke D, Bök P B, et al. Leverage LiFi in Smart Manufacturing [C]// 2020 IEEE Globecom Workshops, December, 7–11, 2020, Taipei, Taiwan. New York: IEEE, 2020: 1–6.
30. Alturjman F, Cacciagrano D, Mostarda L, et al. Light Communication for Controlling Industrial Robots [C]// Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, March, 4–7, 2021, University of Camerino, Italy. Berlin: Springer International Publishing, 2021: 111–128.
31. Sun Zhongyi. Research on network access point scheme of IoT devices based on LiFi [J]. Information Recording Materials, 2020, 21(3): 195–196.
32. Zhang X B, Cao Z Z, Li J H, et al. Wide-Coverage Beam-Steered 40-Gbit/s Non-Line-of-Sight Optical Wireless Connectivity for Industry 4.0 [J]. Journal of Lightwave Technology, 2020, 38(24): 6801–6806.
33. Kouhini S M, Kottke C, Ma Z, et al. LiFi positioning for industry 4.0 [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2021, 27 (6): 1–15.
34. Daniel S, Abhijeet S, Holger F, et al. Impact of industrial environments on visible light communication [J]. Optics Express, 2021, 29(11): 16087–16104.
35. Koonen T, Mekonnen K A, Cao Z, et al. Beam-Steered Optical Wireless Communication for Industry 4.0 [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2021, 27(6): 1–10.
36. Kang Lingchaun. Research on mobile receiving technology of array visible light communication [D]. Zhengzhou: Strategic Support Force Information Engineering University, 2021.
37. Ma Z Y, Kouhini S M, Kottke C, et al. LiFi Positioning and Optimization in an Indoor Factory Environment [C]// IECON 2022-48th Annual Conference of the IEEE Industrial Electronics Society, October, 17–20, 2022, Brussels, Belgium. New York: IEEE, 2022: 1–6.
38. Salvi S, Geetha V. A Novel 2D LED Matrix and Aztec Pattern Inspired Optical Camera Communication for Industrial IoT [C]// 2022 IEEE 19th India Council International Conference, November 24–26, 2022, Kochi, India. New York: IEEE, 2022: 1–6.
39. Kouhini S M. All-Optical Distributed MIMO for LiFi: Spatial Diversity Versus Spatial Multiplexing [J]. IEEE Access, 2022, 10(13): 102646–102658.
40. Costa W, Camporez H, Hinrichs M, et al. Toward AI-Enhanced VLC Systems for Industrial Applications [J]. 2023, 41(4): 1064–107.
41. Reguera V A, Teixeira L, Barriquello C H, et al. Efficient and Low-Complexity Rate and Dimming Control of VLC for Industrial IoT Applications [J]. IEEE Journal of Emerging and Selected Topics in Industrial Electronics, 2022, 3(4): 1087–1095.
42. Xue J S, Ye Z W, Zhang H Y, et al. Flexible Design of Low-Delay MEC-VLC Integrating Network Based on Attocell Overlap for IIoT [J]. Electronics, 2022, 11(6): 223–228.
43. Xue J S, Zhang Y Y, Wang C, et al. Boundaryless Visible Light Communication System Without Attocell for Industrial Internet of Things [J]. IEEE Access, 2022, 10(6): 87454–8746.
44. Xu Xianying. Research on visible light OFDM communication system based on Hartley transform [D]. Dalian: Dalian Maritime University, 2022.
45. Yang Jianguo. Research on semi-blind time synchronization algorithm for visible light communication DCO-OFDM system [D]. Harbin: Harbin Institute of Technology, 2020.
46. Ozyurt A B, Popoola W O. CoMP-JT Scheme for D2D Communication in Industrial LiFi Networks [J]. IEEE Access, 2022, 10(2): 70760–70768.
47. Yang Borui. Research of PLC and VLC cascade communication system based on OFDM [D]. Xi'an: Xi'an Technological University, 2021.
48. Nguyen H, Nguyen V L, Tran D H, et al. Deep Learning for 2D-MIMO scheme based on Optical Camera Communication [C]// 2023 International Conference on Information Networking, January, 11–14, 2023, Bangkok, Thailand. New York: IEEE, 2023: 797–800.

49. Wang Hao. Research on the application of visible light communication technology in intelligent industrial lighting [J]. China New Telecommunications, 2023, 25(2): 13–15.
50. Du X X, Zhang Y Y, Ye Z W, et al. LED lighting area recognition for visible light positioning based on convolutional neural network in the industrial internet of things [J]. Optics Express, 2023, 31(8): 12778–12788.
51. Zhou H Y, Zhang M L, Ren X M. Design and Implementation of Wireless Optical Access System for VLC-IoT Networks [J]. Journal of Lightwave Technology, 2023, 41(8): 2369–2380.
52. Aboagye S, Ndjiongue A R, Ngatched T M N, et al. RIS-Assisted Visible Light Communication Systems: A Tutorial [J]. IEEE Communications Surveys and Tutorials, 2023, 25(1): 251–288.
53. Ozyurt A B, Popoola W O. LiFi-Based D2D Communication in Industrial IoT [J]. IEEE Systems Journal, 2023, 17(1): 1591–1598.
54. Xiao Zhiqiang. Analysis of visible light communication system under fading channel [D]. Nanjing, Nanjing University of Posts and Telecommunications, 2020.
55. Wen Xiangyi, Wang Jingyuan, XU Zhiyong, et al. Analysis and simulation of indoor visible light power distribution [J]. Military communication technology, 2013, 34(01): 73–76.
56. Xia Zhongjin. Research on environmental parameter monitoring in coal mine based on visible light communication [D]. Huainan: Anhui University of Science& Technology, 2021.
57. Wang Y, Tao L, Chi N. High speed WDM VLC system based on multi-band cap64 with weighted pre-equalization and modified CMMA based post-equalization [J]. IEEE communications letters: A publication of the IEEE Communications Society, 2014, 18(10): 1719–1722
58. Xu L D, He W, Li S C. Internet of things in industries: a survey [J]. IEEE Transactions on Industrial Informatics, 2014, 10 (4): 2233–2243.
59. Fusco G, Russo M. Robust MIMO design of decentralized voltage controllers of PV systems in distribution networks [J]. IEEE Transactions on Industrial Electronics, 2017, 46(6): 4610–4620.
60. Striccoli D, Boggia G, Grieco L A. A markov model for characterizing IEEE 802. 15. 4 MAC layer in noisy environments [J]. IEEE Transactions on Industrial Electronics, 2015, 62 (8): 5133–5142.
61. Brundu F G, Pattie, Osello A, et a1. IoT software infrastructure for energy management and simulation in smart cities [J]. IEEE Transactions on Industrial Informatics, 2017, 13 (2): 832–840.
62. Qiao Qi, Zhang Yue. Research on visible light communication system based on MIMO [J]. Wireless Connectivity Technology, 2021, 18(1): 3–4+25.
63. Ke Xizheng, Yang Shangjun, Wu Jiali, et al. Research progress of wireless optical MIMO technology and space-time coding [J]. Journal of Hunan University of Science and Engineering, 2021, 42(05): 25–39.
64. Hang Cheng. Research on channel correlation and interchannel interference in indoor MIMO visible light communication system [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2020.
65. Telatar E. Capacity of Multi-antenna Gaussian Channels [J]. European Transactions on Telecommunications, 1999, 10(6): 585–595.
66. Zhang Jiankun, Yang Yu, Chen Hongda. Modulation Scheme Analysis of Indoor Visible Light Communications [J]. Chinese Journal of Lasers, 2011, 38(4): 0405003.
67. Harald H, Liang Y, Wang Y L, et al. What is LiFi? [J]. Journal of Lightwave Technology, 2016, 34(6): 1533–1544.
68. Ke Xizheng, Lei Sichen, Li Beilei. Experimental Study on Free Space Optical Orthogonal Frequency Division Multiplexing System [J]. Chinese Journal of Lasers, 2015, 42(12): 1205003.
69. Yang Jiawang, CAO Yang, LIU Shitao. Performance research of indoor visible light communication system based on adaptive OFDM technology [J]. Optical Communication Technology, 2015, 39(4): 50–52.
70. Haas H, Cheng C, Brien D O. A guide to wireless networking by light [J]. Progress in Quantum Electronics, 2017, 55(5): 88–111.

71. Li Shaoju. Research on downlink channel and resource scheduling of indoor visible light communication [D]. Chongqing: Chongqing University of Posts and Telecommunications, 2019.
72. Li Jiandong, Sheng Min. Communication network foundation [M]. Beijing: China higher education press, 2004, 8.
73. Liang Jingyuan, Xu Yaxin, Han Meimiao, et al. Research progress of heterogeneous integration system for indoor visible light communication [J]. Optical Communication Technology, 2023, 47(01): 46–50.
74. Qiao Jaijia. Research on Physical layer security of heterogeneous RF/visible light communication networks for multiple users [D]. Guilin: Guilin University of Electronic Technology, 2022.
75. Chen C, Zhong W D, Yang H L, et al. Reduction of SINR Fluctuation in Indoor Multi-Cell VLC Systems Using Optimized Angle Diversity Receiver [J]. Journal of Lightwave Technology, 2018, 36(17): 343–352.
76. Ijaz M, Ghassemlooy Z, Pesek J, et al. Modeling of Fog and Smoke Attenuation in Free Space Optical Communications Link Under Controlled Laboratory Conditions [J]. Journal of Lightwave Technology, 2013, 31(11): 1720–1726.